

PROGRAMME DE FORMATION

Déployer une IA locale et sécurisée pour l'équipe informatique

Installez un assistant IA privé, publiez-le sur le réseau interne et connectez-le à la documentation technique en maîtrisant les accès et les données.

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

Finalité de la formation

Formation de trois jours pour dimensionner, installer et sécuriser une IA locale reliée à la documentation de l'équipe informatique.

Public concerné : Administrateurs systèmes et réseau, architectes infrastructure et responsables techniques

Prérequis : Maîtriser Linux, les réseaux IP, la virtualisation ou les conteneurs et les principes de sécurité d'un service web.

Objectifs opérationnels

- Choisir une architecture et un modèle adaptés au besoin.
- Installer un moteur d'inférence dans un environnement isolé.
- Publier le service avec HTTPS, authentification et filtrage réseau.
- Connecter l'assistant à une documentation technique maîtrisée.
- Superviser, sauvegarder et maintenir la plateforme.

Compétences développées

- Estimer les besoins CPU, mémoire, stockage et GPU.
- Segmenter les composants et limiter les flux réseau.
- Construire une chaîne de recherche documentaire traçable.
- Définir des procédures d'exploitation et de mise à jour.

Programme détaillé

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

Jour 1 - Architecture et installation d'une IA locale

Matin - Besoins, modèles et dimensionnement

- Cas d'usage internes et niveaux de confidentialité.
- Familles de modèles, quantification, fenêtre de contexte et licences.
- Dimensionnement CPU, mémoire, GPU, stockage et concurrence.
- Architecture cible, flux réseau et critères de choix.

Travaux pratiques : Élaborer une matrice de dimensionnement pour trois scénarios d'entreprise.

Après-midi - Mise en service du moteur d'inférence

- Installation en machine virtuelle ou conteneur dans le laboratoire.
- Téléchargement contrôlé d'un modèle et gestion du stockage.
- Premiers appels API, paramètres d'inférence et limites.
- Mesure de la latence, de la mémoire et du débit.

Travaux pratiques : Déployer un moteur local et comparer deux configurations de modèle.

Jour 2 - Publication et sécurisation sur le réseau interne

Matin - Exposition du service

- Interface web, API et séparation des composants.
- DNS interne, reverse proxy, certificats TLS et règles de pare-feu.
- Authentification, groupes, quotas et limitation de débit.
- Gestion des erreurs sans fuite d'information.

Travaux pratiques : Publier le service derrière un reverse proxy HTTPS avec contrôle d'accès.

Après-midi - Données, journaux et durcissement

- Cartographie des données, rétention et politique de journalisation.
- Segmentation réseau, moindre privilège et comptes de service.
- Gestion des modèles, mises à jour et dépendances.
- Tests de sécurité fonctionnels et revue de configuration.

Travaux pratiques : Durcir l'architecture et vérifier les flux effectivement autorisés.

Jour 3 - Assistant documentaire et exploitation

Matin - Recherche dans la documentation interne

- Préparation, découpage et métadonnées des documents.
- Indexation, recherche sémantique et constitution du contexte.
- Citations, droits d'accès et prévention des réponses non sourcées.
- Jeu de tests représentatif et mesure de la qualité.

Travaux pratiques : Indexer un corpus de procédures réseau et produire des réponses avec sources.

Après-midi - Mise en exploitation

- Supervision, métriques, alertes et capacité.
- Sauvegarde des configurations, index et données nécessaires.
- Plan de mise à jour, retour arrière et continuité de service.
- Documentation d'architecture et procédure opérateur.

Travaux pratiques : Finaliser un assistant privé documenté, sécurisé et supervisé.

Projet ou atelier de synthèse

Assistant IA privé accessible sur le réseau interne et capable d'interroger une documentation technique avec citations.

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

- Serveur ou machine virtuelle de laboratoire disposant des ressources adaptées.
- Accès administrateur à l'environnement, DNS de laboratoire et certificats de test.
- Corpus documentaire non sensible fourni pour 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.