

OBJECTIF DE LA FORMATION

À la fin des trois jours, les participants disposeront d'une compréhension complète du cycle de développement d'une application utilisant l'API OpenAI, depuis la formulation des requêtes jusqu'à la sécurisation, l'optimisation et la mise en production.

PARTICIPANT : Toute personne souhaitant intégrer des modèles d'intelligence artificielle dans leurs applications.

PRÉ-REQUIS : Connaître les principes de base du développement logiciel.

FORMATEURS : Spécialiste en développement d'application avec l'IA

MODALITÉS ET PÉDAGOGIE

- Questionnaire d'évaluation en début et en fin de formation
- Cas pratiques et études de cas, Quiz
- Horaires : 9h00 à 12h30 – 13h30 à 17h00
- Nombre maximum de stagiaires : 8
- Formation disponible à distance ou en présentiel

- Méthode interactive et intuitive
- Support de formation remis aux participants
- Assistance Technique et Pédagogique
- Attestation de fin de stage
- Questionnaire de fin de formation

PROGRAMME

Jour 1 — Découvrir et maîtriser les fondamentaux de l'API OpenAI

1. Introduction à l'intelligence artificielle générative

Principes généraux des modèles de langage
Différences entre intelligence artificielle traditionnelle et intelligence artificielle générative
Notions de modèle, contexte, jeton, requête et réponse
Capacités et limites des modèles génératifs
Principaux usages dans les applications professionnelles

2. Présentation de la plateforme OpenAI

Organisation générale de la plateforme
Création et configuration d'un projet
Gestion des clés d'API
Présentation des bibliothèques de développement
Documentation et ressources disponibles

Principes de facturation et de consommation

3. Architecture d'une application utilisant l'API

Communication entre une application cliente, un serveur et l'API OpenAI

Construction d'une requête HTTP
En-têtes, authentification et corps de requête
Format JSON des échanges
Gestion des réponses et des erreurs

Protection des clés d'API côté serveur

4. Découverte de la Responses API

Rôle de la Responses API
Structure d'une requête
Paramètres principaux
Gestion des instructions et des entrées utilisateur
Lecture et exploitation de la réponse
Différences entre la Responses API et les interfaces historiques
Gestion des réponses synchrones et diffusées en continu

5. Sélection et paramétrage des modèles

Familles de modèles disponibles
Modèles généralistes et modèles de raisonnement
Critères de choix d'un modèle
Qualité, rapidité et coût
Taille du contexte
Paramètres influençant la génération
Gestion de la longueur des réponses

6. Conception des instructions

Principes du prompt engineering

Définition claire d'un objectif

Séparation des instructions, du contexte et des données

Utilisation de rôles et de règles

Contraintes de ton, de format et de longueur

Réduction des réponses imprécises ou hors sujet

Techniques de décomposition d'une demande complexe

7. Gestion des erreurs et de la fiabilité

Principaux codes d'erreur

Erreurs d'authentification et de configuration

Limites de requêtes et saturation

Délais d'attente

Stratégies de nouvelle tentative

Journalisation des erreurs

Validation des réponses avant utilisation

Jour 2 — Concevoir des applications enrichies avec l'API OpenAI

1. Gestion des conversations

Construction d'un échange conversationnel

Conservation de l'historique

Gestion du contexte applicatif

Instructions système et messages utilisateur

Limitation de la taille du contexte

Résumé et compression de l'historique

Gestion des sessions et des utilisateurs

2. Génération de contenus textuels

Génération, reformulation et synthèse de textes

Extraction d'informations

Classification de contenus

Traduction et adaptation rédactionnelle

Analyse de documents

Production de réponses adaptées à un métier

Contrôle du style et du niveau de détail

3. Sorties structurées

Intérêt des réponses structurées

Définition d'un format de sortie

Utilisation de schémas JSON

Validation des données produites

Gestion des champs obligatoires et facultatifs

Intégration des réponses dans une base de données ou une interface

Traitement des sorties non conformes

4. Appel de fonctions et utilisation d'outils



Principe de l'appel de fonctions
Description d'une fonction accessible au modèle
Définition des paramètres d'entrée
Détection de l'outil sélectionné par le modèle
Exécution de la fonction par l'application
Transmission du résultat au modèle
Gestion des appels multiples
Contrôle des autorisations et validation des arguments

5. Recherche documentaire et embeddings

Principe des représentations vectorielles
Fonctionnement des embeddings
Similarité sémantique
Indexation de contenus
Recherche d'informations dans une base documentaire

Sélection des passages pertinents
Génération augmentée par la recherche
Gestion de la pertinence et des sources

6. Traitement de contenus multimodaux

Analyse d'images
Association de textes et d'images dans une requête
Extraction d'informations visuelles
Génération et transformation d'images
Principes de traitement des contenus audio
Transcription et génération vocale
Contraintes liées aux formats et à la taille des fichiers

7. Diffusion progressive des réponses

Principe du streaming
Réception progressive des événements
Affichage en temps réel dans une interface
Gestion de l'interruption d'une génération
Traitement des erreurs pendant la diffusion
Amélioration de la perception des performances

Jour 3 — Sécuriser, optimiser et mettre en production une application

1. Architecture d'une application prête pour la production

Séparation entre interface, serveur applicatif et services d'intelligence artificielle
Organisation des composants
Centralisation des appels à l'API
Gestion des environnements de développement, de test et de production
Configuration par variables d'environnement
Gestion des dépendances
Scalabilité et disponibilité

2. Sécurité des applications

Protection et rotation des clés d'API
Gestion des droits d'accès
Authentification des utilisateurs
Validation et filtrage des entrées
Prévention des injections de prompt
Sécurisation des appels de fonctions
Protection des données sensibles
Gestion des journaux et des informations confidentielles

3. Modération et utilisation responsable

Identification des contenus sensibles
Utilisation des mécanismes de modération
Mise en place de règles de filtrage
Gestion des refus et des réponses inappropriées

Supervision humaine
Signalement des contenus problématiques
Information des utilisateurs sur les limites du système
Respect des politiques d'utilisation

4. Maîtrise des coûts

Compréhension de la consommation en jetons
Estimation du coût d'une requête
Suivi de la consommation par application ou utilisateur
Limitation de la taille des entrées et des sorties
Choix du modèle en fonction de la tâche
Réutilisation et mise en cache des contenus
Définition de quotas et de budgets
Mise en place d'alertes de consommation

5. Optimisation des performances

Réduction du temps de réponse
Utilisation du streaming
Exécution asynchrone
Mise en cache des résultats
Parallélisation des traitements
Réduction des appels inutiles
Gestion des limites de débit
Stratégies de reprise en cas d'échec

6. Évaluation de la qualité

Définition de critères de réussite
Constitution d'un jeu de données d'évaluation
Mesure de la pertinence et de la conformité
Comparaison de modèles et de prompts
Identification des hallucinations
Tests de non-régression
Évaluation humaine et automatisée
Suivi de la qualité dans le temps

7. Observabilité et maintenance

Journalisation des requêtes et des réponses
Suivi des erreurs et des temps de réponse
Traçabilité des appels de fonctions
Création d'indicateurs de performance
Analyse des retours utilisateurs
Gestion des versions de prompts
Mise à jour des modèles et des bibliothèques
Maintenance continue de l'application

8. Préparation du déploiement

Vérifications avant mise en production
Gestion des secrets
Tests de charge et de sécurité
Définition des mécanismes de secours
Documentation technique
Procédures d'exploitation
Surveillance après déploiement
Amélioration continue



DOLFI Formation

91 rue Saint Lazare 75009 PARIS – RCS : B438093791 – Siret : 438 093 791 00046

Tel : 01 42 78 13 83 – Fax : 01 40 15 00 53 – formation@dolfi.fr

Retrouver toutes les informations sur nos stages : <https://www.dolfi.fr>



Accès spécifique
consulter notre page

<https://www.dolfi.fr/ressources>