

## Développer des applications d'IA générative sur AWS

*S'approprier les services AWS d'IA Générative pour la réalisation de vos solutions applicatives*

### DESCRIPTION

Vous démarrez la conception d'une application d'IA générative hébergée sur AWS ? Comment démystifier et partager la compréhension autour des Grands modèles de langage (LLMs) pour tous les opérationnels impliqués ? Comment aider les data scientist à s'approprier les services d'IA générative proposés par AWS, notamment Bedrock ?

Cette formation s'adresse à des équipes techniques intervenant dans la réalisation de produits digitaux faisant appel à l'IA générative.

Ce module de 2 jours fournit un aperçu de l'IA générative et présente les fondamentaux de l'utilisation des modèles de langage (LLM) et s'appuie de plusieurs démonstrations permettant de s'en approprier les bonnes pratiques. L'accent est mis sur la conception et l'implémentation d'une application d'IA générative dans un environnement AWS, notamment grâce à Amazon Bedrock, LangChain et le prompt engineering.

### OBJECTIFS PEDAGOGIQUES

- Découvrir le concept d'IA générative, son importance, ses avantages et ses risques potentiels
- Identifier la valeur métier et les cas d'usage de l'IA générative
- Définir les fondements techniques et la terminologie clé de l'IA générative
- Décrire les concepts, avantages, cas d'usage, l'architecture et la structure de coûts d'Amazon Bedrock
- Mettre en oeuvre les meilleures pratiques et techniques de prompt engineering pour les modèles fondation (FM)
- Inventorier les composants, personnaliser les modèles fondation (FM) et tirer parti des services AWS pour les applications d'IA générative
- Intégrer LangChain avec les LLMs, les prompts, les chaînes et les agents pour Amazon Bedrock

### PUBLIC CIBLE

- Développeur
- Data Engineer
- Tech Lead
- Data Scientist

### PRE-REQUIS

#### Stage pratique

IA

Code :

**AWSIA**

Durée :

**2 jour(s) (14,00 heures)**

Exposés : **70 %**

Cas pratiques : **25 %**

Echanges d'expérience : **5 %**

#### Inter-entreprises :

Prochaines sessions disponibles [sur notre site web](#).

Tarif : 1 595,00 € HT /

participant

#### Intra-entreprise :

Tarifs et dates sur demande.

- Connaissances de base du Cloud AWS (niveau AWS Technical Essentials)
- Niveau intermédiaire en Python

#### **METHODE PEDAGOGIQUE**

Formation avec apports théoriques, échanges sur les contextes des participants et retours d'expérience pratique des formateurs, complétés de travaux pratiques et de mises en situation.

#### **PROFIL DES INTERVENANTS**

Cette formation est dispensée par un-e ou plusieurs consultant-es d'OCTO Technology ou de son réseau de partenaires, expert-es reconnus des sujets traités.

Le processus de sélection de nos formateurs et formatrices est exigeant et repose sur une évaluation rigoureuse leurs capacités techniques, de leur expérience professionnelle et de leurs compétences pédagogiques.

#### **MODALITÉS D'ÉVALUATION ET FORMALISATION À L'ISSUE DE LA FORMATION**

L'évaluation des acquis se fait tout au long de la session au travers des ateliers et des mises en pratique. Afin de valider les compétences acquises lors de la formation, un formulaire d'auto-positionnement est envoyé en amont et en aval de celle-ci. Une évaluation à chaud est également effectuée en fin de session pour mesurer la satisfaction des stagiaires et un certificat de réalisation leur est adressé individuellement.

#### **PROGRAMME PEDAGOGIQUE DETAILLE**

##### **Jour 1**

##### **OUVERTURE DE SESSION**

- Accueil des stagiaires et tour de tables des attentes
- Présentation du programme des deux jours de formation

##### **INTRODUCTION À L'IA GÉNÉRATIVE - L'ART DU POSSIBLE**

- Panorama du Machine Learning
- Principes de bases de l'IA générative
- Cas d'usage de l'IA générative
- L'IA générative en pratique
- Opportunités et risques

### **PLANIFICATION D'UN PROJET D'IA GÉNÉRATIVE**

- Fondamentaux de l'IA générative
- L'IA générative appliquée
- Contexte et enjeux
- Les étapes de planification d'un projet d'IA générative
- Identifier les risques et mettre en place des actions correctives

### **PREMIERS PAS AVEC AMAZON BEDROCK**

- Présentation d'Amazon Bedrock
- Architecture et cas d'usage
- Comment utiliser Amazon Bedrock
- Démonstration : configuration de l'accès à Amazon Bedrock et utilisation des "playgrounds"

### **FONDEMENTS DU PROMPT ENGINEERING**

- Bases des modèles fondamentaux
- Principes clés du prompt engineering
- Techniques de prompt de base
- Techniques avancées de prompt
- Démonstration : affiner un prompt texte simple
- Techniques spécifiques à chaque modèle
- Identifier et corriger les usages inappropriés des prompts
- Réduction des biais
- Démonstration : atténuer les biais dans la génération d'images

## **Jour 2**

### **COMPOSANTS DES APPLICATIONS AMAZON BEDROCK**

- Panorama des applications et cas d'usage
- Principaux composants des applications d'IA générative
- Modèles fondamentaux et interface FM
- Exploitation des jeux de données et des embeddings
- Démonstration : création et utilisation de word embeddings
- Composants applicatifs complémentaires
- RAG (Retrieval-Augmented Generation)
- Ajustement (fine-tuning) des modèles
- Sécurisation des applications d'IA générative
- Architecture type d'une application d'IA générative

### **LES MODÈLES FONDAMENTAUX D'AMAZON BEDROCK**

- Présentation des modèles fondamentaux disponibles dans Amazon Bedrock
- Utiliser les FMs pour les inférences
- Méthodes et bonnes pratiques d'Amazon Bedrock
- Protection des données et traçabilité
- Mise en pratique : générer du texte avec un prompt zero-shot grâce à un modèle Bedrock

### **INTRODUCTION À LANGCHAIN**

- Optimiser les performances des LLM
- Intégrer AWS avec LangChain
- Exploiter les modèles via LangChain
- Construire des prompts efficaces
- Organisation des documents avec des indexes
- Stockage et récupération des données avec la mémoire
- Utiliser les "chains" pour séquencer les composants
- Gestion des ressources externes avec les agents LangChain

### **SCHÉMAS D'ARCHITECTURE**

- Introduction aux modèles d'architecture
- Résumé automatique de texte
- Mise en pratique : résumer de courts fichiers avec Amazon Titan Text Premier
- Mise en pratique : résumer des longs textes avec Amazon Titan
- Question/Réponse automatisée
- Mise en pratique : utiliser Amazon Bedrock pour mettre en place un système de Q&A
- Chatbots
- Mise en pratique : construire un chatbot
- Génération de code
- Mise en pratique : utiliser les modèles BedRock pour générer du code
- LangChain et agents intégrés avec Amazon BedRock
- Mise en pratique : créer des applications conversationnelles avec l'API Converse

---

#### **Accessibilité**

L'inclusion est sujet important pour OCTO Academy.

---

**OCTO Academy est une marque d'OCTO Technology, SAS au Capital de 509525.30 €**

34 avenue de l'Opéra, 75002 Paris - 01 70 36 18 67 - academy@octo.com - <https://www.octo.academy>

SIRET 418 166 096 00069 – RCS PARIS B 418 166 096

Organisme de formation certifié Qualiopi - Déclaration d'activité N°11 754893675 auprès du préfet de la Région d'Ile-de-France  
Cet enregistrement ne vaut pas agrément de l'Etat.

---

Nos référent-es sont à votre disposition pour faciliter l'adaptation de votre formation à vos besoins spécifiques.  
Pour les contacter : [academy.accessibilite@octo.com](mailto:academy.accessibilite@octo.com)