

Mis à jour le 26/10/2026

S'inscrire

Formation Arize Phoenix

2 jours (14 heures)

Présentation

Arize Phoenix permet de comprendre, évaluer et améliorer les applications reposant sur les grands modèles de langage. Avec OpenInference et OpenTelemetry pour instrumenter les traitements, les traces et les évaluations deviennent des leviers pour fiabiliser vos RAG et vos agents IA en conditions réelles.

Notre formation Arize Phoenix vous permettra de dépasser les vérifications ponctuelles pour construire une démarche structurée d'observabilité LLM, du diagnostic local au suivi en production.

Vous apprendrez à instrumenter une application, à explorer ses traces et à identifier les défaillances de recherche documentaire, de génération ou d'appel d'outils. Vous évaluerez automatiquement les réponses et calibrez un LLM juge à partir d'annotations humaines et de critères métier explicites.

À l'issue de la formation, vous serez en mesure de comparer des versions de prompts ou de modèles avec des expériences reproductibles pour décider sur des résultats mesurables. Vous saurez interpréter les compromis entre qualité, latence et coûts.

Cette formation aborde également la protection des données, le monitoring et le rôle d'Arize AX dans les opérations de production. Quatre ateliers vous guideront dans une boucle complète d'instrumentation, de diagnostic, d'évaluation et d'amélioration continue.

Comme toutes nos formations, celle-ci vous présentera **la dernière version stable** de la technologie et ses nouveautés.

Objectifs

- Expliquer les principes de l'observabilité LLM et les rôles respectifs de Phoenix et d'Arize AX
- Instrumenter une application avec OpenInference et exploiter ses traces dans Phoenix
- Diagnostiquer les défaillances d'un RAG et d'un agent utilisant des outils
- Automatiser les évaluations et calibrer un LLM juge sur des annotations humaines
- Comparer des versions de prompts ou de modèles par des expériences reproductibles
- Définir un suivi de production intégrant sécurité, qualité, latence et coûts

Public visé

- Ingénieur IA
- Développeur d'applications LLM
- Data Scientist
- Ingénieur MLOps
- Architecte logiciel

Pré-requis

- Pratique de la programmation avec Python et de la gestion des dépendances
- Connaissance des principes des LLM et des techniques de prompting
- Première expérience d'un RAG ou d'un agent utilisant des outils
- Capacité à consommer une API REST et à manipuler des données au format JSON dans une application

Pré-requis techniques

- Ordinateur sous Windows 11 ou macOS ou Linux avec 8 Go de RAM minimum, 16 Go recommandés pour les ateliers
- Environnement Python 3.11 avec pip et droits d'installation des dépendances pédagogiques
- Éditeur Visual Studio Code ou environnement JupyterLab et navigateur récent
- Connexion Internet stable autorisant les téléchargements de paquets et les appels aux services de modèles utilisés
- Accès à une instance Arize Phoenix locale ou fournie pour la session et clé API d'un fournisseur de modèles avec quota disponible

Programme de notre formation Arize Phoenix

[Jour 1 - Matin]

Comprendre et instrumenter l'observabilité LLM

- Enjeux de l'observabilité LLM : comportement, qualité, erreurs, latence et consommation de tokens
- Rôles de Phoenix et d'Arize AX : diagnostic, évaluation, expérimentation et exploitation en production
- Modèle de traçage : traces, spans, relations parent-enfant, attributs et propagation du contexte

- Articulation entre OpenInference et OpenTelemetry : conventions sémantiques, instrumentation et export des traces
- Installation de Phoenix, configuration des projets et instrumentation automatique ou manuelle d'une application Python
- Atelier pratique : Instrumenter une application LLM avec OpenInference, envoyer ses traces à Phoenix et vérifier les attributs collectés.

[Jour 1 - Après-midi]

Tracer et diagnostiquer un RAG et un agent

- Lecture d'une trace RAG : requête, recherche documentaire, contexte transmis et réponse générée
- Diagnostic des défauts de retrieval : documents non pertinents, contexte incomplet et sources obsolètes
- Analyse de la génération : instructions, fidélité au contexte, erreurs et latence par étape
- Traçage d'un agent : orchestration, appels d'outils, arguments, résultats et trajectoire d'exécution
- Investigation des échecs agentiques : mauvais outil, boucle, erreur d'exécution et consommation excessive de tokens
- Atelier pratique : Diagnostiquer un RAG et un agent défaillants, localiser les étapes responsables et vérifier une correction dans les traces.

[Jour 2 - Matin]

Évaluer les réponses et calibrer un LLM juge

- Construction d'un jeu d'évaluation représentatif : exemples issus des traces, cas limites et réponses de référence
- Choix des critères de qualité : pertinence documentaire, exactitude, fidélité au contexte et réussite de tâche
- Combinaison des évaluateurs déterministes, des annotations humaines et du LLM juge
- Conception d'une grille de jugement : consignes, labels, exemples et justification des scores
- Calibration du juge : accord avec les annotations humaines, analyse des désaccords, biais et validation sur des exemples distincts
- Atelier pratique : Annoter un échantillon, automatiser son évaluation dans Phoenix et ajuster le LLM juge en analysant ses désaccords avec les évaluations humaines.

[Jour 2 - Après-midi]

Expérimenter et préparer le suivi en production

- Organisation des expériences Phoenix : versions de datasets, prompts, modèles et paramètres contrôlés
- Comparaison des variantes sur les mêmes entrées : scores, analyse par exemple, latence, coûts estimés et régressions
- Sécurité des traces : masquage avant export, exclusion des secrets, contrôle des accès et politique de conservation

- Définition du monitoring de production : qualité, erreurs, latence, tokens, coûts et échantillonnage des évaluations
- Exploitation avec Arize AX : tableaux de bord, moniteurs, alertes et réutilisation de l'instrumentation OpenInference
- Atelier pratique : Comparer deux variantes de prompt ou de modèle, retenir une configuration selon les résultats, vérifier le masquage des données et définir les seuils de surveillance en production.

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.