Évaluation, Benchmarking et Tests de Fiabilité des Agents IA en Production

Lien : https://innov-systems.com/formation/evaluation-benchmarking-et-tests-de-fiabilite-des-agents-ia-en-production

DURÉE
5 jours (35h)

RÉFÉRENCE
BDI127

CATÉGORIE
IA Agentique et Systèmes Multi-Agents

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre les enjeux de l’évaluation des agents IA
- Mettre en place des benchmarks fiables et pertinents
- Tester la robustesse des systèmes IA en production
- Mesurer la performance des agents de manière objective
- Identifier les limites et faiblesses des modèles
- Améliorer la fiabilité des systèmes multi-agents
- Utiliser des outils d’évaluation et de monitoring
- Mettre en place des tests automatisés
- Analyser les résultats de performance
- Optimiser les agents IA en continu
- Assurer la qualité des systèmes en production
- Étudier des cas concrets d’évaluation IA

POUR QUI ?

- Ingénieurs IA
- Machine learning engineers
- Data scientists
- MLOps engineers
- Architectes IA
- Développeurs logiciels
- Consultants en intelligence artificielle
- Chefs de projet IA
- Product managers IA
- Responsables qualité logicielle
- Analystes data
- Toute personne impliquée dans le déploiement et la fiabilité des systèmes IA

Innov Systems

Programme détaillé

1 / Introduction à l’évaluation des agents IA

- Importance de l’évaluation en environnement de production
- Différence entre test en développement et en production
- Enjeux de fiabilité et de robustesse

2 / Méthodologies de benchmarking des agents IA

- Définition des critères de performance
- Comparaison de modèles et d’agents
- Construction de jeux de tests représentatifs

3 / Mesure de la performance des agents

- Précision, pertinence et cohérence des réponses
- Latence et efficacité d’exécution
- Mesures qualitatives et quantitatives

4 / Tests de fiabilité et robustesse

- Tests de stress et cas limites
- Gestion des erreurs et comportements inattendus
- Résilience des systèmes multi-agents

5 / Évaluation en conditions réelles (production)

- Monitoring des agents en production
- Analyse des logs et traces d’exécution
- Détection des dérives de performance

Innov Systems

6 / Outils et frameworks d’évaluation

- Solutions de testing automatisé des agents IA
- Outils de benchmark et d’observabilité
- Intégration dans les pipelines MLOps

7 / Amélioration continue des agents IA

- Boucles de feedback utilisateur
- Réentraînement et ajustement des prompts
- Optimisation des workflows agentiques

8 / Sécurité et conformité dans les tests

- Respect des normes et réglementations
- Détection des biais et comportements risqués
- Audit des systèmes IA

9 / Simulation et scénarios de test avancés

- Simulation d’utilisateurs et de charges
- Scénarios multi-agents complexes
- Tests de cohérence inter-agents

10 / Études de cas et projets pratiques

- Mise en place d’un système de test complet
- Analyse de performances réelles d’agents IA
- Bonnes pratiques industrielles

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

12 au 18 Jan. 2027	Présentiel - Casablanca
08 au 12 Mars 2027	Présentiel - Casablanca
03 au 07 Mai 2027	Présentiel - Casablanca
28 Juin au 02 Juil. 2027	Présentiel - Casablanca
23 au 27 Août 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés