

Mis à jour le 10/09/2026

S'inscrire

Formation RAGFlow

Présentation

RAGFlow est une plateforme open source de Retrieval-Augmented Generation permettant de transformer des documents complexes en assistants IA fiables, explicables et adaptés aux usages métier.

Notre formation RAGFlow vous permettra de maîtriser les étapes essentielles de conception d'une solution RAG, depuis l'ingestion des documents jusqu'à la génération de réponses contextualisées avec leurs sources.

Vous apprendrez à configurer les bases de connaissances, à choisir les méthodes de découpage adaptées aux contenus, à sélectionner les modèles de langage et d'embeddings, puis à contrôler la qualité de l'indexation documentaire.

La formation vous aidera également à régler la recherche hybride, les seuils de similarité, le reranking et les stratégies de récupération afin d'améliorer la pertinence des réponses tout en limitant les hallucinations.

Enfin, vous concevrez des assistants conversationnels et des agents exploitant vos connaissances métier. Vous aborderez les bonnes pratiques de sécurité, de supervision, de maintenance et de mise en production d'une solution RAGFlow.

Objectifs

- Comprendre l'architecture et les cas d'usage d'une solution RAG avec RAGFlow
- Créer et alimenter des bases de connaissances adaptées aux documents métier
- Configurer le découpage, les embeddings et l'indexation pour optimiser la recherche
- Régler la récupération hybride, le reranking et les paramètres de pertinence
- Concevoir un assistant conversationnel et un agent reliés à des sources documentaires
- Préparer l'exploitation, la sécurisation et l'évolution d'une plateforme RAGFlow

Public visé

- Développeur IA
- Data Engineer
- Architecte solutions
- Ingénieur MLOps
- Chef de projet IA

Pré-requis

- Connaissance des principes des LLM et de l'IA générative
- Pratique de l'administration d'environnements Linux ou de conteneurs
- Notions de recherche documentaire, de données non structurées et d'API
- Expérience de base avec Python ou un langage de développement équivalent

Pré-requis techniques

- Instance RAGFlow opérationnelle, déployée en local ou sur un environnement partagé
- Docker 24.0.0 minimum et Docker Compose 2.26.1 minimum pour un déploiement autonome
- Machine x86 disposant d'au moins 4 coeurs, 16 Go de mémoire vive et 50 Go d'espace disque pour l'auto-hébergement
- Accès configuré à un fournisseur de LLM et à un modèle d'embeddings

Programme de notre formation RAGFlow

[Jour 1 - Matin]

Fondamentaux du RAG et prise en main de RAGFlow

- Positionnement de RAGFlow dans l'écosystème de l'IA générative et cas d'usage métier
- Principes du Retrieval-Augmented Generation : ingestion, indexation, récupération et génération
- Architecture de la plateforme : services, stockage documentaire, moteur de recherche et modèles IA
- Découverte de l'interface : espaces de travail, jeux de données, assistants, agents et paramètres de modèles
- Déploiement avec Docker Compose, vérification des services et configuration initiale des fournisseurs de modèles

[Jour 1 - Après-midi]

Constitution et préparation des bases de connaissances

- Création d'une base de connaissances et organisation des corpus par domaine métier

- Import de documents structurés et non structurés : PDF, présentations, tableaux, documents bureautiques et contenus texte
- Choix des modèles d'embeddings selon la langue, le corpus et les contraintes de qualité
- Sélection des méthodes de chunking : général, manuel, livre, tableau, présentation et autres modèles documentaires
- Contrôle du parsing, édition des chunks, ajout de mots-clés et gestion des documents actifs
- Atelier pratique : créer une base de connaissances, importer un corpus métier et valider la qualité du découpage documentaire

[Jour 2 - Matin]

Optimisation de la recherche et de la pertinence

- Fonctionnement de la recherche hybride : recherche plein texte, recherche vectorielle et fusion des résultats
- Réglage du seuil de similarité, du poids vectoriel et du nombre de passages récupérés
- Usage du reranking pour améliorer l'ordre des sources et arbitrage entre latence et précision
- Stratégies d'indexation : texte traité, questions générées, contexte enrichi et poids des noms de fichiers
- Découpage hiérarchique, stratégie parent-enfant et enrichissement contextuel des passages retrouvés

[Jour 2 - Après-midi]

Tests de récupération et assistant conversationnel

- Conception de scénarios de test fondés sur les questions réelles des utilisateurs
- Analyse des sources citées, des chunks remontés et des causes de réponses imprécises
- Ajustement itératif du parsing, des mots-clés, des paramètres de recherche et du reranking
- Création d'un assistant conversationnel connecté à une ou plusieurs bases de connaissances
- Rédaction des instructions système : périmètre métier, style de réponse, gestion de l'incertitude et citations
- Atelier pratique : configurer un assistant documentaire, mesurer la qualité de récupération et optimiser ses réponses sur un jeu de questions métier

[Jour 3 - Matin]

Agents et pipelines de données

- Différences entre assistant RAG, agent et workflow orienté traitement de données
- Conception d'un agent : instructions, mémoire, récupération de connaissances et contrôle des sorties
- Découverte des composants de pipeline : parsing, chunking, enrichissement, indexation et routage
- Automatisation de l'ingestion avec les pipelines de données et synchronisation des sources externes
- Introduction aux intégrations par API et MCP pour connecter RAGFlow aux applications métier

Industrialisation, gouvernance et mise en production

- Architecture cible : dimensionnement, choix du moteur documentaire, persistance et configuration des services
- Sécurisation des accès aux modèles, aux documents et aux données sensibles
- Gestion du cycle de vie des connaissances : mise à jour, désactivation, synchronisation et traçabilité
- Supervision des assistants et des agents : journaux d'exécution, analyse des erreurs et indicateurs de qualité
- Bonnes pratiques de déploiement : séparation des environnements, sauvegarde, évolution des modèles et validation avant production
- Atelier pratique : finaliser une solution RAGFlow de bout en bout, définir son plan de mise en production et présenter les critères de validation

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.