

**LSI**LEARNY SOCIETY
INTERNATIONAL

PLAQUETTE 2026

IA & TECHNOLOGIES · ASSISTANTS, AGENTS & AUTOMATISATION

Créer une base de connaissances IA pour son entreprise

Réunir, trier et structurer les documents de l'entreprise pour que l'IA y trouve la bonne réponse, la cite et ne l'invente pas.

● Cours disponible · sur devis

DURÉE

2 jours (14h)

FORMAT

Présentiel · Distanciel · Hybride

PUBLIC

Responsables de la documentation, du support, de la qualité, des RH ou de la formation, chefs de projet et référents IA chargés de rendre les connaissances de l'entreprise accessibles aux équipes.

PRÉREQUIS

Utiliser régulièrement un outil d'IA générative. Disposer d'un accès, autorisé par l'entreprise, à un outil capable d'interroger des documents (NotebookLM, projet Claude, GPT avec fichiers, Copilot sur SharePoint...). Apporter si possible un corpus de 10 à 30 documents sur un même domaine, anonymisé et validé pour cet usage ; à défaut, un corpus fictif est fourni. Aucune compétence en programmation n'est requise.

Pourquoi cette formation.

Deux jours pour construire une base de connaissances que les équipes interrogent en langage courant : procédures, offres, documentation, règles internes. Le travail porte d'abord sur les documents – ce qui entre, dans quel état, avec quels droits – puis sur l'outil, les tests et la mise à jour.

Charger un dossier partagé dans un outil d'IA prend quelques minutes ; obtenir des réponses justes demande autre chose. Deux versions d'une même procédure, un scan illisible, un tableau mal découpé, un brouillon jamais validé : l'outil ne fait pas la différence et répond avec la même assurance. Sans règles d'accès ni de mise à jour, la base vieillit, se contredit et finit par montrer ce qu'elle ne devait pas montrer.

Le premier jour explique simplement comment un outil retrouve un passage – découpage des documents, recherche par le sens, principe du RAG : chercher d'abord, rédiger ensuite à partir de ce qui a été trouvé – et en tire des règles de préparation du corpus. Le second jour, chaque participant construit sa base dans l'outil autorisé (NotebookLM, projet Claude, GPT avec fichiers, Copilot sur SharePoint...), la soumet à une batterie de questions de test, contrôle les citations et fixe sa gouvernance : propriétaire, droits, versions, rythme de revue.

OBJECTIFS

À l'issue de la formation.

01 Expliquer simplement comment un outil retrouve un passage dans des documents et s'en sert pour répondre (recherche par le sens, RAG)

02 Évaluer la qualité d'un corpus : sources, versions, doublons, contradictions, données sensibles

03 Structurer des documents pour qu'ils soient retrouvés et compris par l'IA

04 Construire une base de connaissances interrogeable dans l'outil autorisé par l'entreprise

05 Rédiger des instructions de réponse qui imposent la citation des sources et le refus de répondre hors du corpus

06 Tester la base avec une batterie de questions et analyser les erreurs obtenues

07 Organiser les droits d'accès, le versioning et la mise à jour de la base

Le programme.

JOUR 1 • MODULE 01 • 1 H 15

Interroger ses documents plutôt que la mémoire de l'IA

- Interroger l'IA « de mémoire » ou sur ses documents : deux usages, deux niveaux de fiabilité
- Le RAG en une phrase : l'outil cherche d'abord les passages utiles dans vos documents, puis rédige sa réponse à partir d'eux
- Les formes courantes : carnet de sources, projet avec fichiers, assistant avec fichiers, copilote branché sur l'espace documentaire
- Ce que la base change — réponses citées, périmètre maîtrisé — et ce qu'elle ne change pas : un document faux donne une réponse fausse

JOUR 1 • MODULE 02 • 1 H 30

Comment l'outil retrouve un passage : découpage, recherche et embeddings

- Le découpage des documents en passages, et ce qu'il casse : tableaux, listes, renvois « voir ci-dessus »
- Recherche par mots-clés et recherche par le sens : ce que chacune trouve (référence exacte, formulation différente)
- Les embeddings sans jargon : chaque passage reçoit des coordonnées de sens, et l'outil ramène les passages les plus proches de la question
- Proche ne veut pas dire juste : un passage obsolète ou hors sujet peut être ramené avec les autres
- L'outil ne relit pas tout le corpus à chaque question : pourquoi une information présente peut rester introuvable

JOUR 1 • MODULE 03 • 2 H

Cartographier et auditer ses sources

- Définir le périmètre : quelles questions, pour quels utilisateurs, avec quels droits
- Inventaire : où sont les documents, qui les rédige, qui les valide
- Qualité d'une source : auteur, date, statut (validé, brouillon, obsolète), propriétaire
- Doublons, versions concurrentes, contradictions : une information, une source de référence
- Ce qui n'entre pas : données personnelles, documents confidentiels, documents hors du périmètre de droits

Structurer des documents lisibles par l'IA

- Un document, un sujet ; des titres qui disent ce que contient la section
- En-tête normalisé : titre, version, date, propriétaire, public visé
- Convertir ce que l'outil lit mal : tableaux en image, scans, diaporamas, schémas
- Glossaire des acronymes et du vocabulaire interne
- Fiches questions-réponses pour les demandes les plus fréquentes, et conventions de nommage des fichiers

Construire la base dans l'outil autorisé

- Choisir la forme selon le besoin et l'outil : carnet de sources, projet, GPT avec fichiers, Copilot sur l'espace documentaire
- Instructions de réponse : s'en tenir aux documents, citer document et section, répondre « non documenté » sinon
- Signaler les contradictions entre documents plutôt que d'en choisir une
- Charger par lots et contrôler que chaque document est réellement exploité
- Préparer les questions d'exemple proposées aux utilisateurs

Citations, hallucinations et tests

- Exiger la citation, puis vérifier qu'elle dit bien ce que la réponse affirme
- Les erreurs typiques d'une base : invention, versions mélangées, passage sorti de son contexte, trou comblé de mémoire
- Construire une batterie de questions de test avec réponses attendues et sources
- Questions hors périmètre et questions pièges : ce que la base doit refuser
- Corriger à la bonne place : le document, sa structure ou les instructions
- Fixer un seuil d'acceptation avant d'ouvrir la base aux utilisateurs

Permissions, versioning et mise à jour

- Qui peut interroger la base peut en lire tout le contenu : une base par périmètre de droits
- Outils branchés sur un espace documentaire : ils s'appuient en principe sur les droits existants, dont les erreurs deviennent visibles — à vérifier avec l'administrateur
- Versioning : remplacer un document, ne jamais empiler les versions ; tenir un journal des modifications
- Propriétaire de la base, référents par domaine, rythme de revue, date de péremption des documents
- Faire vivre la base : recueillir les questions restées sans réponse pour compléter le corpus

Projet final : une base interrogeable, testée et gouvernée

- Assembler corpus audité, documents structurés, instructions, tests et gouvernance
- Passer la batterie de tests finale et en documenter les résultats
- Démonstration devant le groupe et réponses à des questions imprévues
- Plan de lancement : annonce aux utilisateurs, questions d'exemple, première revue datée

Comment on travaille.

PÉDAGOGIE

- Explications sans jargon, appuyées sur des démonstrations en direct et une activité débranchée pour comprendre la recherche documentaire
- Fil rouge sur deux jours : de l'audit du corpus à la base testée, sur les documents de chaque participant ou sur un corpus fictif fourni
- Plus de la moitié du temps en pratique : ateliers en sous-groupes, exercices individuels, tests réels dans l'outil
- Analyse critique des réponses obtenues : citations contrôlées, erreurs classées, corrections apportées au corpus
- Démonstration finale devant le groupe et feedback du formateur sur la base et sa gouvernance
- Chaque exercice se fait avec l'outil autorisé dans l'entreprise ; à défaut, avec les versions gratuites de ChatGPT, Claude ou Copilot et des cas fictifs fournis par le formateur.

ÉVALUATION

Diagnostic avant la formation, livrables contrôlés et quiz en cours de parcours, base finale testée et évaluée sur grille. Attestation de formation remise à l'issue du parcours.

VOUS REPARTEZ AVEC

- Votre base de connaissances interrogable, testée sur votre corpus ou sur le corpus fourni
- La grille d'audit des sources : statut, version, propriétaire, données sensibles
- Le guide de structuration des documents pour l'IA : titres, en-têtes, glossaire, questions-réponses
- Le modèle d'instructions de réponse : s'en tenir aux documents, citer, dire « non documenté »
- La batterie de questions de test et la grille de classement des erreurs
- La fiche de gouvernance : propriétaire, droits, versioning, rythme de mise à jour

INTRA, INTER OU SUR MESURE

Construisons votre session.

TARIF

Sur devis, communiqué avant tout engagement

RÉPONSE

Sous 48 à 72 h ouvrées

MODALITÉS

Présentiel, classe virtuelle ou hybride

CONTACT

contact@lsi-society.com · 06 69 53 48 71

Demander un devis → lsi-society.com/catalogue/base-connaissances-ia

LSI — Learnly Society International — Entreprise individuelle Marcus Détrez, 135 avenue du 24 avril 1915, Bâtiment E2, 13012 Marseille, France · SIRET 897 775 094 00039. Déclaration d'activité enregistrée sous le n° 93132113413 auprès du préfet de région Provence-Alpes-Côte d'Azur. Cet enregistrement ne vaut pas agrément de l'État. Exonération de TVA en application de l'article 261-4-4°-a du code général des impôts (formation professionnelle continue).