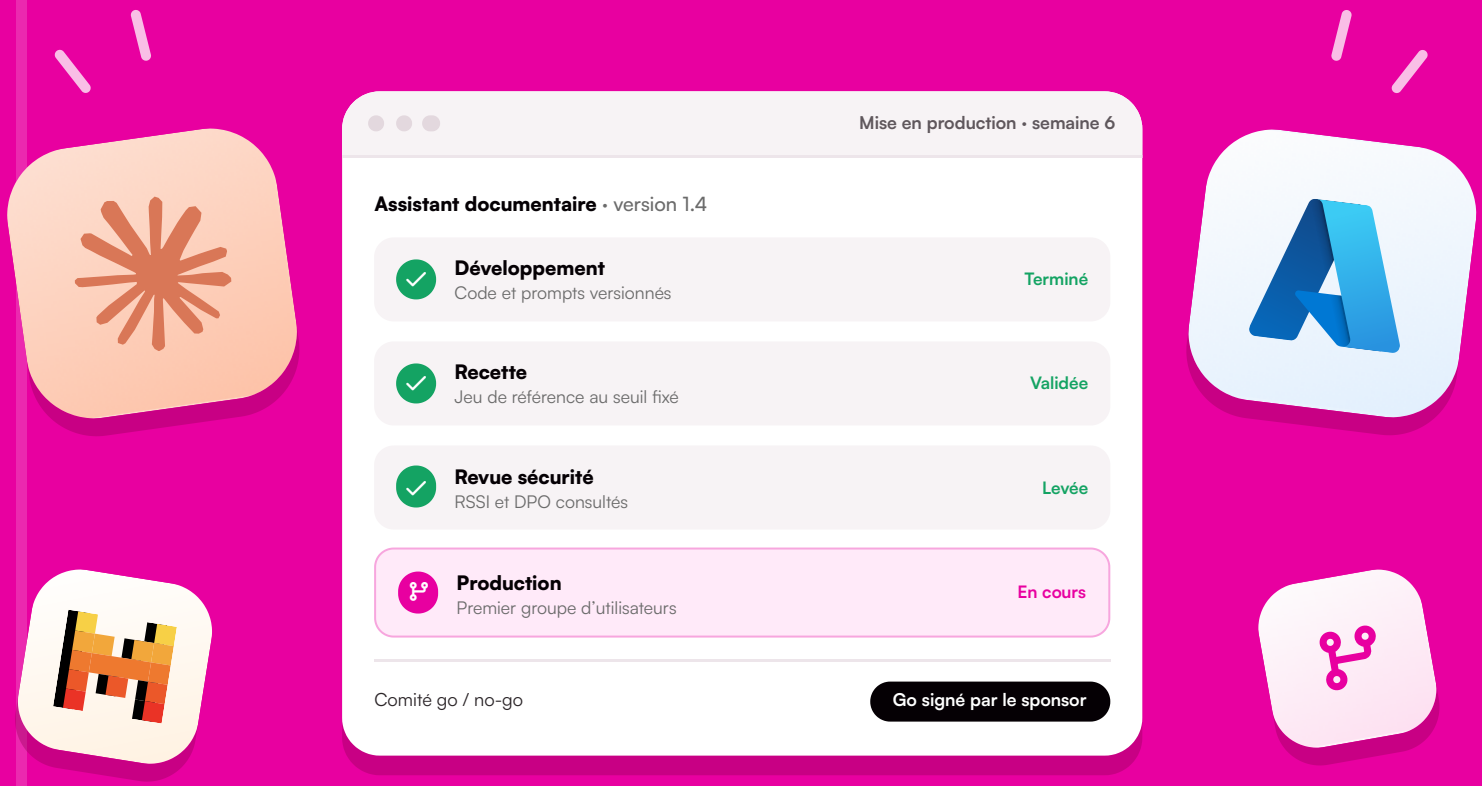
AI sisters

Du POC à la production en 6 semaines

le chemin de mise en production d'un projet IA, tel qu'on le pratique au Lab



Guide DSI | 3 environnements | 1 schéma d'architecture réutilisable

OFFICIAL PARTNER Claude Partner Network

AI Sisters est partenaire officiel du Claude Partner Network, le réseau de partenaires d'Anthropic.

AI sisters

Du POC à la production en 6 semaines

Le guide DSI pour faire sortir un projet IA du prototype

Un POC qui impressionne en démo et qui ne sert jamais, la plupart des DSI en ont déjà vu passer. Le prototype tourne, mais personne n'a prévu où il vivra, comment on le teste ni qui le maintient.

Ce guide décrit la méthode du Lab AI Sisters : six semaines, un cas d'usage, trois environnements, une recette écrite dès la première semaine et un comité qui décide du passage en production.

Vous y trouverez le calendrier, le schéma d'architecture que nous reprenons d'un projet à l'autre, les garde-fous, ce qu'on surveille une fois en ligne et la liste de contrôle avant le go.

OFFICIAL PARTNER Claude Partner Network

AI Sisters est partenaire officiel du Claude Partner Network, le réseau de partenaires d'Anthropic.

1. Pourquoi les POC restent des POC

Quatre causes reviennent presque à chaque fois. On les traite dès la semaine 1, pas à la fin.

Ce qu'on observe	Ce qu'on fait dès la semaine 1
Le prototype vit sur un poste	On ouvre les trois environnements le premier jour, avec la DSI. Le POC est construit là où il tournera.
« Ça marche » n'est pas défini	Le métier écrit un jeu de cas réels et les seuils d'acceptation avant qu'on voie le moindre résultat.
La sécurité découvre le projet en fin de course	Le RSSI et le DPO voient l'architecture en semaine 1 et la version candidate en semaine 5.
Personne ne sait qui le maintient	Le kit de run et le contact d'astreinte font partie des livrables, au même titre que le code.

Les huit parties de ce guide

1 Pourquoi les POC restent des POC 2
2 Le calendrier des six semaines 3
3 Les trois environnements 4
4 Le schéma d'architecture 5
5 La recette 6
6 Les garde-fous 7
7 L'observabilité et le run 8
8 Le comité go / no-go 9

2. Le calendrier des six semaines

Un seul cas d'usage. Chaque fonctionnalité suit le même cycle : deux semaines de développement, une restitution avec le métier, deux séries d'itérations de deux jours, puis une restitution de validation.

Semaine 1 **Cadrage**	On fixe le cas d'usage et ce qu'on laisse de côté. Le métier écrit les cas de test et les seuils. On choisit l'architecture et on demande les accès : données, comptes, SSO, environnements. Ces demandes sont souvent le vrai chemin critique, d'où leur place au premier jour.
Semaines 2 et 3 **Construction**	Développement sur l'environnement de développement, avec les données de test validées. Point hebdomadaire avec l'équipe projet. Restitution en fin de semaine 3 devant le métier, suivie d'un compte rendu écrit.
Semaine 4 **Itérations**	Deux séries de deux jours sur les retours du métier. Le jeu de référence tourne à chaque changement de prompt, de modèle ou de données. Les cas pièges entrent dans la recette.
Semaine 5 **Conditions réelles**	Les utilisateurs pilotes travaillent sur l'environnement de recette avec des données réelles. Revue sécurité de la version candidate. Démo au sponsor.
Semaine 6 **Mise en production**	Plus aucune nouvelle fonctionnalité : corrections et réglages seulement. Comité go / no-go, bascule en production pour un premier groupe d'utilisateurs, remise du kit de run et passation à vos équipes.

Nos trois règles. Jamais deux semaines sans restitution. La démo a lieu avant la dernière semaine. La dernière semaine sert aux itérations, jamais à ajouter une fonctionnalité.

Côté AI Sisters

Ingénieur IA senior	Conçoit l'architecture, développe, tient la recette à jour
Chef de projet IA	Pilote le rythme, anime les restitutions, rédige les comptes rendus

Côté client

Sponsor métier	Valide chaque jalon et signe le go
Référent DSI	Ouvre les accès et les environnements
RSSI et DPO	Revue en semaines 1 et 5
Utilisateurs pilotes	Testent en semaine 5, donnent leurs retours

3. Les trois environnements

Ils existent dès la semaine 1. Une version ne monte d'un cran que si elle a passé les tests du cran précédent.

Développement

On construit

Chez AI Sisters

- Données de test ou anonymisées
- Accès réservé à l'équipe projet
- On peut tout casser, c'est fait pour

Recette

On vérifie

Hébergée par AI Sisters, accessible en continu par vos équipes

- Même configuration que la production
- Données réelles ou représentatives, validées par le DPO
- Utilisateurs pilotes et jeu de référence

Production

On sert

Chez vous (Azure, AWS, cloud souverain) ou hébergée par AI Sisters

- Connexion par votre SSO
- Droits d'accès existants respectés
- Aucune modification directe

Ce qui passe d'un environnement à l'autre, et comment

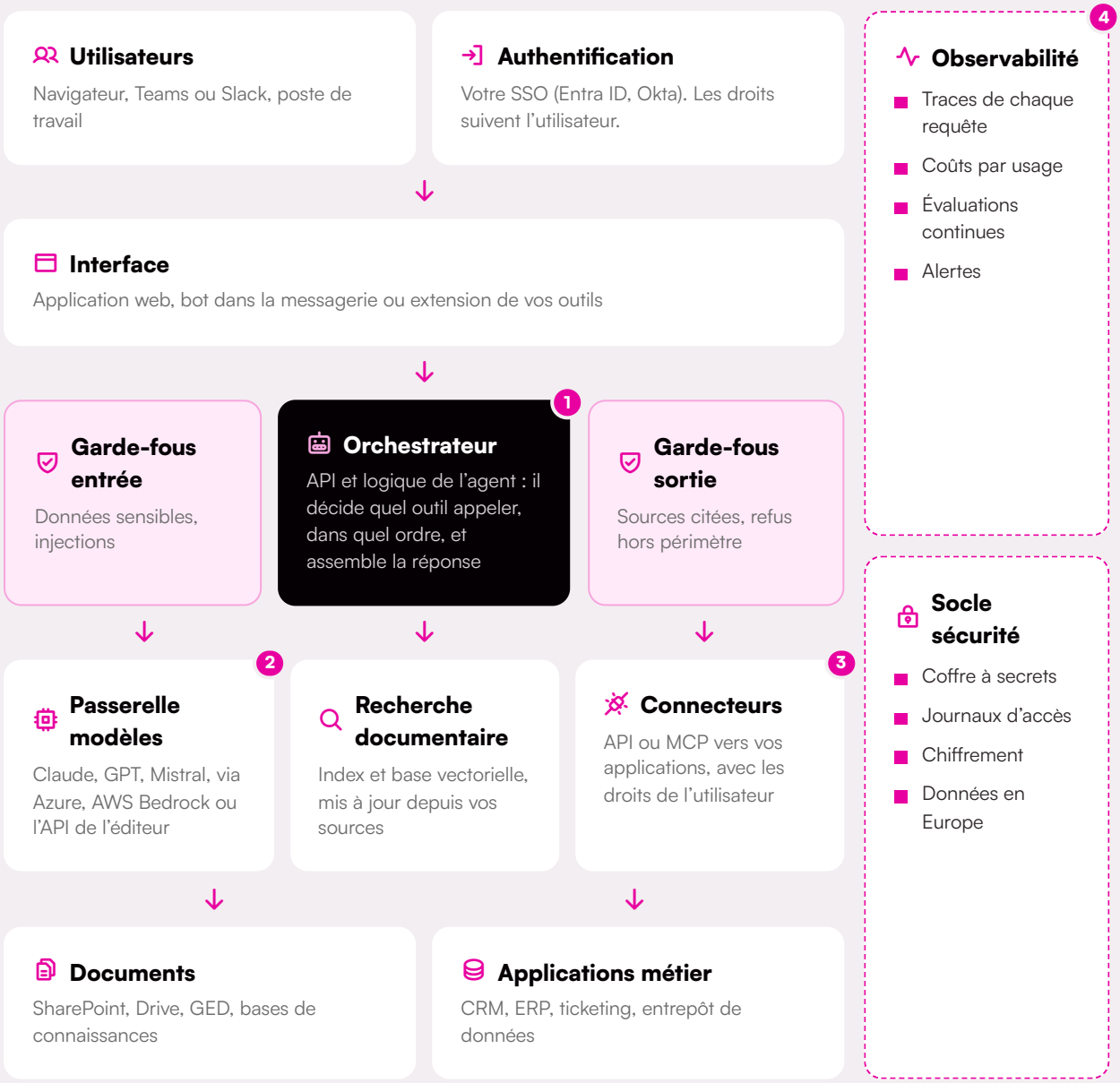
Le code	Versionné dans un dépôt Git, déployé par la même chaîne d'intégration sur les trois environnements.
Les prompts et les réglages du modèle	Versionnés comme du code. On ne retouche jamais un prompt à la main en production.
Les secrets	Clés d'API et mots de passe dans un coffre-fort (Azure Key Vault, AWS Secrets Manager), jamais dans le code.
Les données	Elles ne descendent jamais de la production vers le développement sans anonymisation.
L'infrastructure	Décrite en code (Terraform, Bicep) pour que la production soit une copie conforme de la recette.

À compléter : vos environnements

Hébergement		**Région**	
Modèle		**SSO**	
Dépôt		**Réf. DSI**	

4. Le schéma d'architecture

Celui que nous reprenons d'un projet à l'autre, qu'il s'agisse d'un assistant documentaire, d'un agent métier ou d'un outil d'analyse. Les briques changent de fournisseur, la structure reste.



1. L'orchestrateur ne touche jamais une source directement. Il passe par les connecteurs ou la recherche documentaire.
2. La passerelle permet de changer de modèle ou de fournisseur sans réécrire l'application.
3. Les connecteurs sont en lecture seule par défaut. Toute écriture passe par une validation humaine.
4. Chaque brique envoie ses traces au même endroit : c'est ce qui rend le run possible.

Comment s'en servir. Reprenez ce schéma dans votre dossier d'architecture, remplacez chaque brique par l'outil retenu chez vous, et faites-le relire par le RSSI dès la semaine 1. C'est le document qui servira de base à la revue sécurité de la semaine 5.

5. La recette

Un modèle de langage ne répond pas deux fois exactement pareil. On ne teste donc pas une réponse, on teste un taux de réponses conformes sur un jeu de cas réels, à chaque changement.

Le test	Ce qu'on vérifie	Quand
Jeu de référence	Les réponses sur 30 à 100 cas réels écrits par le métier, avec la réponse attendue et la source	À chaque changement de prompt, de modèle ou de données
Non-régression	Rien de ce qui marchait la veille ne casse	Avant chaque montée en recette
Cas pièges	Questions hors périmètre, information absente, tentatives de contournement des consignes	Semaines 4 et 5
Recette utilisateurs	L'outil sert-il vraiment ? Temps gagné, clarté, confiance dans les réponses	Semaine 5, utilisateurs pilotes
Charge et coût	Temps de réponse et coût par requête au volume attendu	Semaine 5
Droits d'accès	Un utilisateur ne voit que ce qu'il a le droit de voir, testé avec un compte non administrateur	Semaine 5

Notre conseil Fixez les seuils en semaine 1, avant d'avoir vu les premiers résultats. Sinon on finit toujours par trouver que 70 % « c'est déjà pas mal ».

À compléter : vos critères d'acceptation

Réponses conformes visées _____

Temps de réponse maximum _____

Coût maximum par requête _____

Nombre de cas du jeu de référence _____

Qui signe la recette _____

6. Les garde-fous

Un garde-fou à chaque endroit où l'outil peut mal faire. Ils sont dans le schéma dès la semaine 1, pas ajoutés après l'incident.

Où	Le garde-fou	Exemple
Accès	SSO et droits hérités : l'outil ne voit que ce que voit la personne qui l'utilise.	Un commercial ne peut pas interroger les contrats RH.
Entrée	Détection des données sensibles et des tentatives d'injection avant l'envoi au modèle.	Un IBAN collé dans la question est masqué.
Actions	Lecture seule par défaut. Validation humaine pour tout ce qui écrit, envoie, supprime ou paie.	L'agent prépare le mail, la personne l'envoie.
Sortie	Sources citées, refus clair hors périmètre, format de réponse contrôlé.	« Je ne trouve pas cette information dans les documents. »
Modèle	Offre entreprise sans entraînement sur vos données, hébergement en Europe si nécessaire.	Contrat et DPA relus par le DPO.
Coûts	Plafond de consommation par utilisateur et par jour, alerte sur le budget mensuel.	Une boucle d'agent ne vide pas le budget en une nuit.

La règle

Aucun agent n'écrit dans un système de production sans validation humaine tant que la recette n'a pas montré qu'il peut s'en passer. Et cette décision se prend en comité, pas dans le code.

Les questions à poser au RSSI et au DPO en semaine 1

01 Quelles données ont le droit de sortir du SI, et vers quel fournisseur de modèle ?

02 Le traitement doit-il être ajouté au registre RGPD ? Faut-il une analyse d'impact ?

03 Quelles applications l'agent pourra-t-il lire, et lesquelles lui restent fermées ?

7. L'observabilité

Un outil IA peut se dégrader sans tomber en panne : un document change, un modèle est mis à jour, les questions évoluent. On ne le voit que si on trace tout.

Ce qu'on enregistre pour chaque requête

- Utilisateur pseudonymisé
- Question
- Documents retrouvés
- Outils appelés
- Réponse
- Modèle et version du prompt
- Temps de réponse
- Coût
- Avis de l'utilisateur

Le tableau de bord	Ce qui déclenche une alerte
Usage	Chute brutale du nombre d'utilisateurs actifs d'une semaine à l'autre
Qualité	Score du jeu de référence sous le seuil, pic d'avis négatifs
Performance	Temps de réponse ou taux d'erreur au-dessus du seuil
Coûts	Budget du mois consommé plus vite que prévu

7. Le run : qui appelle qui

La question que tout DSI pose, à juste titre : et si ça plante le vendredi soir ?

Le kit de run, remis en semaine 6

- Documentation technique
- Guide utilisateur
- Procédure de maintenance
- Procédure de retour arrière

Le retour arrière

La version précédente du prompt, du modèle et du code reste prête à redéployer. On la teste au moins une fois avant le go.

À compléter : votre organisation de run

Support niveau 1 _____

Support niveau 2 _____

Astreinte et horaires _____

Qui décide d'un retour arrière _____

8. Le comité go / no-go

Dix cases à cocher en semaine 6. Le sponsor métier, le référent DSI et le RSSI sont dans la salle. Une case vide, c'est un no-go ou un go avec réserves écrites.

- [ ] Jeu de référence au-dessus du seuil fixé en semaine 1
- [ ] Non-régression au vert sur la version candidate
- [ ] Revue sécurité faite, points bloquants levés
- [ ] Avis du DPO obtenu, registre de traitement à jour
- [ ] SSO branché, droits testés avec un compte non administrateur
- [ ] Garde-fous testés sur les cas pièges
- [ ] Tableaux de bord et alertes actifs
- [ ] Retour arrière testé au moins une fois
- [ ] Kit de run remis et relu par l'équipe qui reprend
- [ ] Go signé par le sponsor métier

Production ne veut pas dire tout le monde

En semaine 6, on ouvre à un premier groupe d'utilisateurs réels, sur des données réelles. L'élargissement se fait ensuite par vagues, à chaque fois après un coup d'œil aux tableaux de bord.

À compléter : le procès-verbal du comité

Date du comité		**Participants**	
Décision		**Réserves**	
Premier groupe		**Revue à 30 jours**	

Méthode AI Sisters Lab, version de novembre 2026. Les fournisseurs et outils cités le sont à titre d'exemple : le schéma s'adapte à votre environnement.

AI sisters

Un projet IA à faire passer en production ?

Le Lab AI Sisters cadre, développe et met en production des agents et des outils IA avec les DSI et les directions métier. Vos équipes reprennent la main à chaque livraison.

Parler de votre projet

Site

aisisters.ai

E-mail

contact@aisisters.ai

Téléphone

07 80 96 07 26

Adresse

255 boulevard Pereire, 75017 Paris

Cabinet de conseil en IA, partenaire officiel du Claude Partner Network.



Ils ont choisi AI Sisters

ENGIE JCDecaux ALL ACCOR CMA CGM VINCI FACILITIES SAMSUNG HAVAS IPSEN Innovation for patient care

+ de 60 000
collaborateurs accompagnés

+ de 900
entreprises accompagnées de A à Z

9,5/10
de note moyenne donnée à nos formations