Les 30 questions à poser avant de choisir son architecture IA

Six questions par axe, à remplir pour un cas d'usage précis. Elles ne se convertissent pas en score : elles disent ce qu'il faudra vérifier, chiffrer ou écrire dans le contrat.

Cas d'usage : _____ Date : _____

01 Coût total de possession

- [ ] Combien de sièges IA à horizon trois ans ?
  Moins de 500 / 500 à 5 000 / au-dessus de 5 000
- [ ] Volume mensuel de tokens attendu à horizon trois ans ?
  Ordre de grandeur
- [ ] Budget annuel IA disponible sur trois ans ?
  Fourchette
- [ ] Capacité d'investissement initial pour une construction ?
  Oui / Non / Partiel
- [ ] Modèle de facturation acceptable ?
  Par siège / par usage / forfait / mix
- [ ] Coûts cachés identifiés dans votre plan ?
  Formation, maintenance, migration, support

02 Gouvernance

- [ ] Politiques d'usage par département requises ?
  Oui / Non / Selon département
- [ ] Traçabilité par prompt et par utilisateur requise ?
  Standard / Renforcée / Critique
- [ ] Exigences AI Act identifiées et documentées ?
  Oui / En cours / Non
- [ ] Certifications requises ?
  SOC 2, ISO 27001, HIPAA, autres
- [ ] Résidence des données imposée ?
  Europe / France / Selon type de données / Non
- [ ] Politique de rétention des logs et prompts définie ?
  Oui / En cours / Non

03 Agilité

- [ ] Nombre de modèles différents à supporter à horizon deux ans ?
  Un / Deux à trois / Plusieurs
- [ ] Fréquence attendue de changement de modèle sur un cas d'usage ?
  Rare / Régulière / Continue
- [ ] Quel besoin de tester rapidement un nouveau modèle sur un périmètre de production limité ?
  Nécessaire / Souhaitable / Pas prioritaire
- [ ] Fine-tuning maison prévu à horizon deux ans ?
  Oui / Peut-être / Non
- [ ] Cas d'usage spécifiques à un métier différenciant ?
  Un ou plusieurs / En prospection / Aucun
- [ ] Roadmap fournisseur acceptable ?
  Confiance totale / Vigilance / Insuffisante

04 Résilience

- [ ] Tolérance à une indisponibilité fournisseur ?
  Heures / Journée / Jours ouvrés
- [ ] Fallback multi-fournisseur nécessaire en production ?
  Oui / Souhaitable / Non
- [ ] SLA contractuel requis pour la production ?
  Nécessaire / Souhaitable / Non
- [ ] Quel niveau de portabilité des agents et prompts exigez-vous à trois ans ?
  Critique / Importante / Secondaire
- [ ] Plan de sortie fournisseur documenté ?
  Oui / En cours / Non
- [ ] Continuité en cas de rupture contractuelle ?
  Assurée / À construire / Non traitée

05 Expérience utilisateur

- [ ] Intégration bureautique native attendue par les utilisateurs ?
  Critique / Importante / Non prioritaire
- [ ] Effort UX interne disponible pour construire une interface ?
  Équipe dédiée / Ponctuel / Aucun
- [ ] Volume d'utilisateurs à onboarder rapidement ?
  Moins de 500 / 500 à 5 000 / au-dessus de 5 000
- [ ] Temps de formation acceptable par utilisateur ?
  Trente minutes / Une journée / Plus
- [ ] Support utilisateur interne ou externalisé ?
  Interne / Externalisé / Mix
- [ ] Adoption métier prioritaire ou progressive ?
  Prioritaire immédiat / Progressif / Non critique

→ Points restés ouverts

Les questions sans réponse établie deviennent les vérifications à mener avant de décider, chacune avec la personne chargée d'y répondre.

VÉRIFICATION À MENER	QUI RÉPOND	ÉCHÉANCE

> La décision se prend avec la matrice du volet 5 : contraintes éliminatoires d'abord, priorité ensuite. Méthode complète : **mauricemendy.com/la-matrice-de-decision** · Diagnostic en ligne : **diagnostic.mauricemendy.com**