

JOURNÉE
FRANÇAISE
DES TESTS
LOGICIELS



Nabil
Idhammou

David
Jedat

Marius
Ortega

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

9 juin 2026

BEFFROI DE MONTROUGE

onepoint.
beyond the obvious

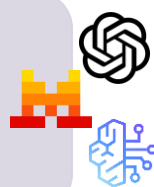
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

Introduction

Les voies d'accès se multiplient


Socles IA d'entreprise

- Azure OpenAI / Azure AI Foundry
- Amazon Bedrock
- Google Vertex AI
- Mistral AI




Assistants bureautiques

- Microsoft 365 Copilot
- Google Gemini Workspace



Assistants code

- GitHub Copilot
- Cursor
- Claude Code
- Codex




Assistants intégrés

- Atlassian Rovo
- Postman AI
- Notion AI
- Figma AI



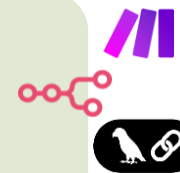
Agents internes

- Agents RH
- Agents IT / support
- Agents test / qualité
- Agents conformité / reporting

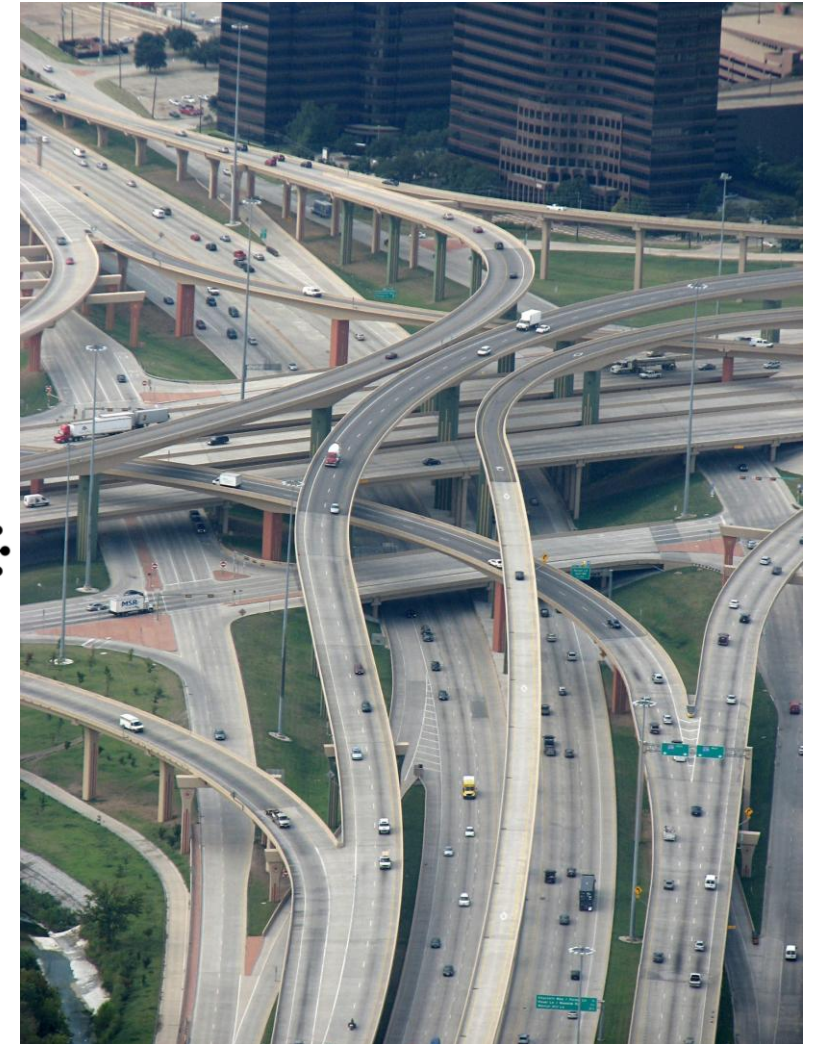


Orchestration, connaissance, gouvernance

- n8n
- Make
- LangGraph / LangChain
- Logic app



Et les modèles économiques évoluent



Qui tient le volant ?

Coût

Combien ça va consommer ?

Qualité

Est-ce fiable ?

Sécurité

Quelles données, quels accès, quelles actions ?

Performance

Est-ce viable à l'échelle ?

Responsabilité

Qui décide, qui arrête, qui assume ?



Sans responsable identifié, les métriques ne permettent pas de décision.

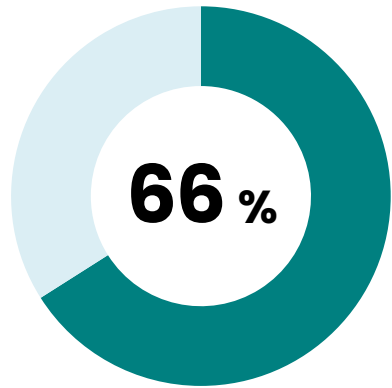
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

Comment passer d'une IA en libre circulation à une IA pilotée, mesurée, utile et viable, sans casser la dynamique d'innovation ?

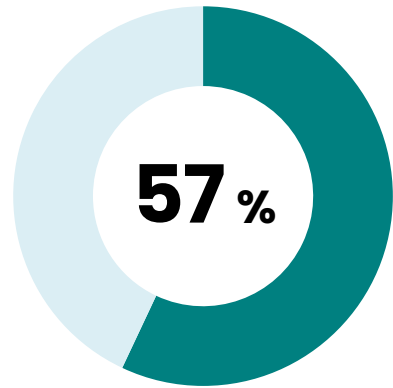
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

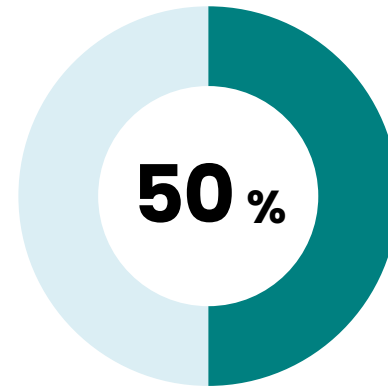
Absence de signalisation ?



De collaborateurs
utilisent l'IA **sans**
vérifier l'**exactitude**
des résultats



déclarent **cache** leur
usage de l'IA au travail



reconnaissent des
usages **non conformes**
aux règles internes

Usages dispersés

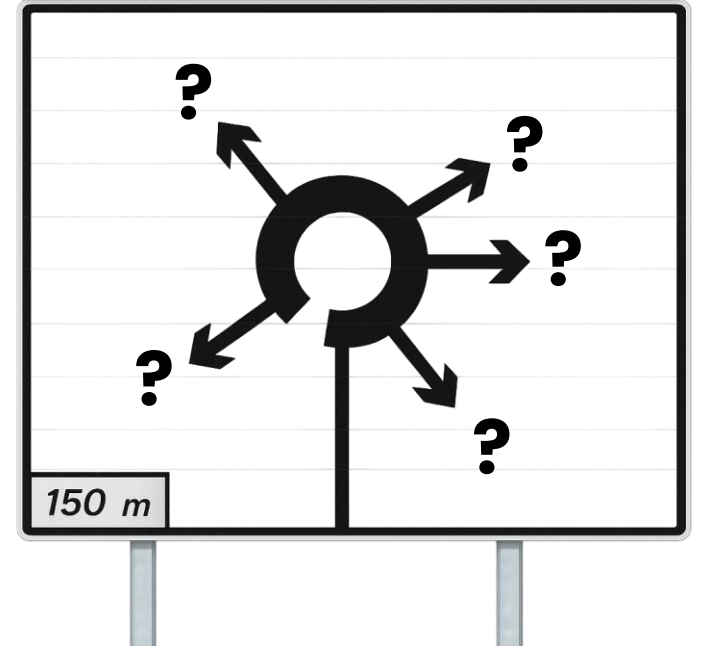
Comptes personnels

Outils non validés

Peu de règles communes

Peu de formation

Responsabilités floues



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

J'accélère Réponse A , je ralentis réponse B

**Un système classique dérive surtout quand on le modifie.
Un système IA peut dériver même si on ne touche pas au code.**

Il doit donc être pilotée en continu.

Piloter \neq Ralentir

Trop cadré

on **ralentit** les essais
on **décourage** les équipes

Pas assez cadré

on multiplie les **usages jetables**
on masque les **coûts** et les **risques**
on découvre **trop tard** les effets d'échelle

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

Péage dans 300m

Rovo s'intègre.
L'usage paraît gratuit.

Accès libre
Les utilisateurs s'habituent.
Les comportements changent.

Puis arrivent les crédits,
les plafonds dépassés, la vraie
consommation facturée.

Jira Standard : **25 crédits / utilisateur / mois**

Rovo Chat : 10 crédits par requête

Rovo Agents : 10 crédits par requête

Deep Research : 100 crédits par requête

1 requête Chat / jour / personne = 20 000 crédits / mois

Quand les coûts arrivent : on arrête ? on bloque ?

On laisse grimper les coûts ou on pilote ?



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

Tourner en rond

Le problème du gaspillage

SORTIES GÉNÉRÉES EN MASSE

CONTENUS REDONDANTS

DOCUMENTS PEU RELUS

RÉSULTATS CORRIGÉS PLUSIEURS FOIS

AGENTS QUI TOURNENT SANS VALEUR DÉMONTRÉE

DOUBLONS DE SOLUTIONS IA ENTRE ÉQUIPES



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

Quand la seule route est bloquée

Une solution IA qui fait tout peut simplifier l'accès aux collaborateurs, mais peut rendre difficile le pilotage

DÉPENDANCE FOURNISSEUR

PROPAGATION DES ERREURS

DIFFICULTÉ À TRAITER TOUS LES USAGES

DÉPASSEMENTS BUDGÉTAIRES DIFFICILES À ATTRIBUER

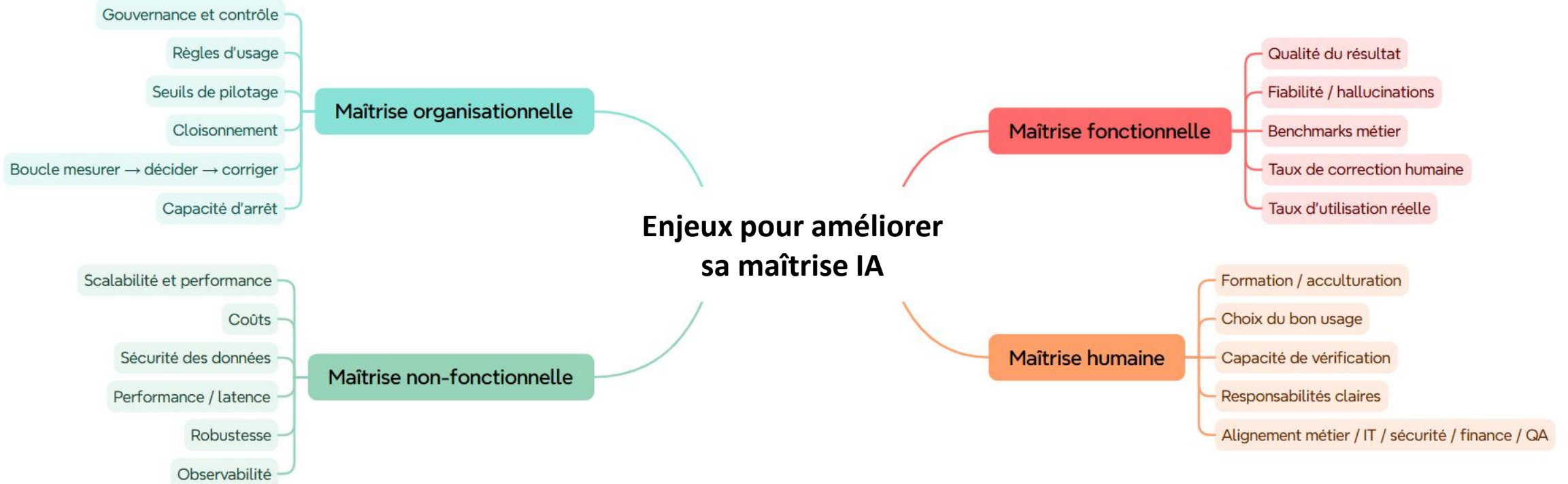
FORTE CRITICITÉ EN CAS D'INCIDENT

Centraliser simplifie l'usage et la visibilité, mais peut concentrer la fragilité. Il peut être important de cloisonner les usages.



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Le banc d'essai avant l'autoroute

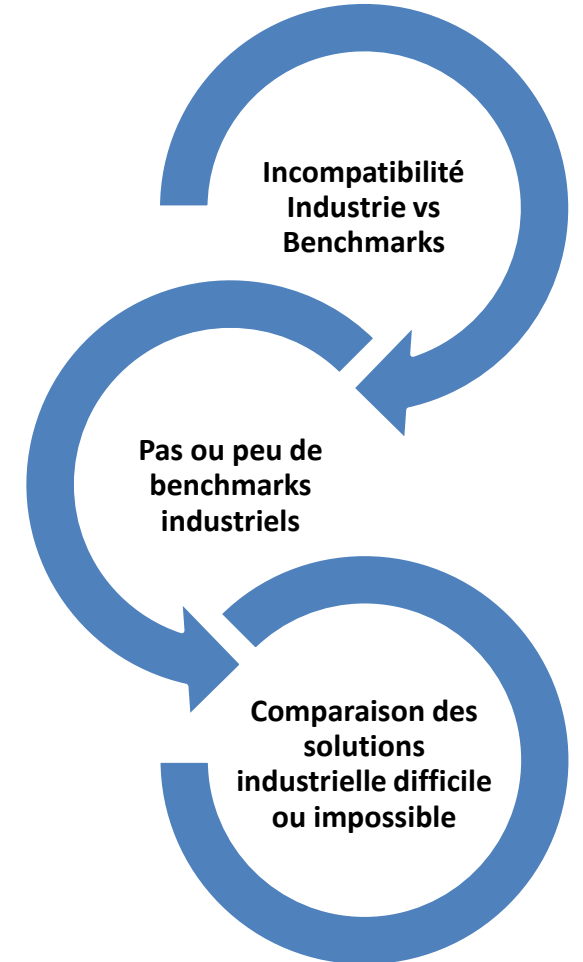
Model	Organization	Global Average	Reasoning Average	Coding Average	Agentic Coding Average
GPT-5 High	OpenAI	79.33	98.17	77.10	46.67
GPT-5 Medium	OpenAI	78.85	96.58	75.05	50.00
GPT-5 Pro	OpenAI	78.73	96.67	72.11	43.33
Claude Sonnet 4.5 Thinking	Anthropic	78.26	95.28	80.36	50.00

LiveBench (White et al)



Les industriels devraient se constituer des benchmarks propriétaires (ou open source si possible) pour évaluer les nouvelles solutions au regard de leurs besoins réels.

Un modèle d'IA ça périmé, pas un benchmark aligné au besoin



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Les quatre voyants à surveiller

On ne pilote pas l'IA avec une seule métrique. On observe les coûts, la qualité, la sécurité, la performance et on y ajoute l'observabilité

OBSERVABILITÉ



Une IA utile mais trop chère, instable, risquée ou avec trop de latence n'est pas « en maîtrise »

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Le prix à la pompe

Le coût du modèle n'est pas le coût du système.

- INFRASTRUCTURE CLOUD
- STOCKAGE DES DONNÉES
- PIPELINES D'INTÉGRATION
- SÉCURITÉ ET GESTION DES ACCÈS
- OBSERVABILITÉ ET LOGS
- TESTS ET BENCHMARKS
- SUPERVISION HUMAINE
- MAINTENANCE DES PROMPTS / AGENTS
- SUPPORT ET GOUVERNANCE



Requêtes

Tokens

Crédits

Actions
agentiques

Quotas

Limites /
pay-as-you-go

Le prix d'un appel IA ne reflète pas le coût réel d'une IA en production. Le coût réel inclut l'infrastructure, la sécurité, l'observabilité, les tests, les données, les pipelines, les outils et la maintenance. **Ils se cumulent et évoluent**

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Le prix à la pompe

Le coût du système IA ne suffit pas pour calculer un ROI

COÛT PAR TÂCHE

COÛT PAR LIVRABLE VALIDÉ

COÛT DE CORRECTION HUMAINE

COÛT DES RÉGÉNÉRATIONS

COÛT D'OBSERVABILITÉ

COÛT D'EXPLOITATION



Le coût à mesurer concerne aussi celui des résultats et de leur valeur apportée ou non.

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

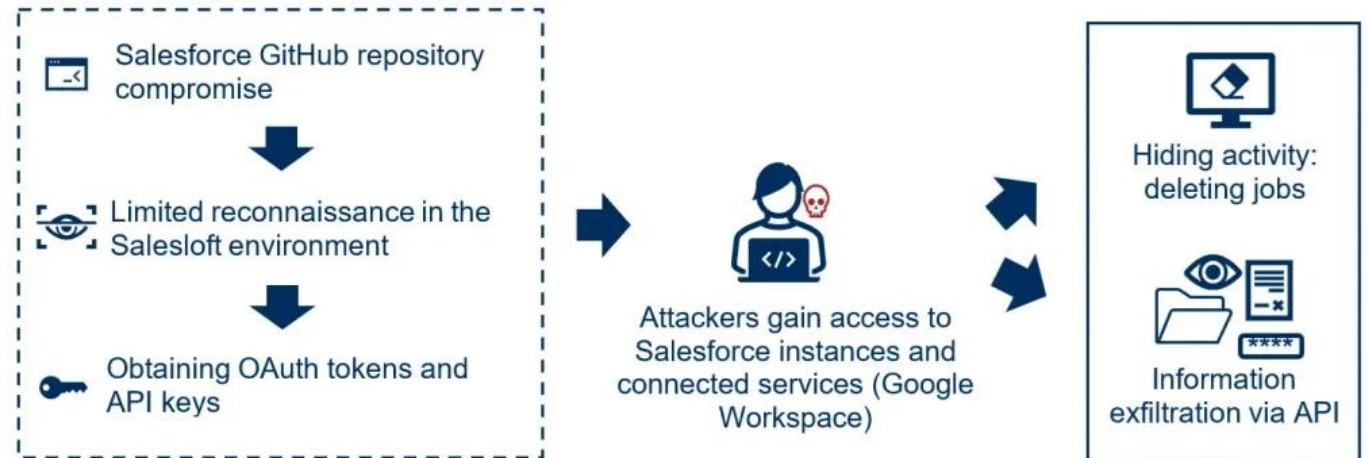
L'importance des glissières

DRIFT[®]
a Salesloft company



Août 2025 : des **tokens OAuth compromis** liés à Salesloft Drift (addon Salesforce) ont permis d'accéder à des données Salesforce de plusieurs entreprises.

- **Outil touché** : Salesloft Drift connecté à Salesforce
- **Point compromis** : tokens OAuth / refresh tokens
- **Clients publics touchés** : Cloudflare, Palo Alto Networks, Zscaler, Workday
- **Données exposées** : secrets, contacts, comptes, opportunités, tickets support, informations commerciales



Un add-on IA ou agentique connecté au SI devient une nouvelle voie d'accès.

La maîtrise doit couvrir le modèle, mais aussi les connecteurs, les droits, les tokens, les logs et la capacité d'arrêt.

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

L'importance des glissières

À contrôler :

données accessibles

données interdites

outils appelables

API utilisables

actions autorisées

actions interdites

décisions sous
validation humaine



Niveaux d'autonomie IA :

faible risque : automatisation possible

risque moyen : validation humaine

risque fort : contrôle strict

risque critique : supervision renforcée ou
interdiction

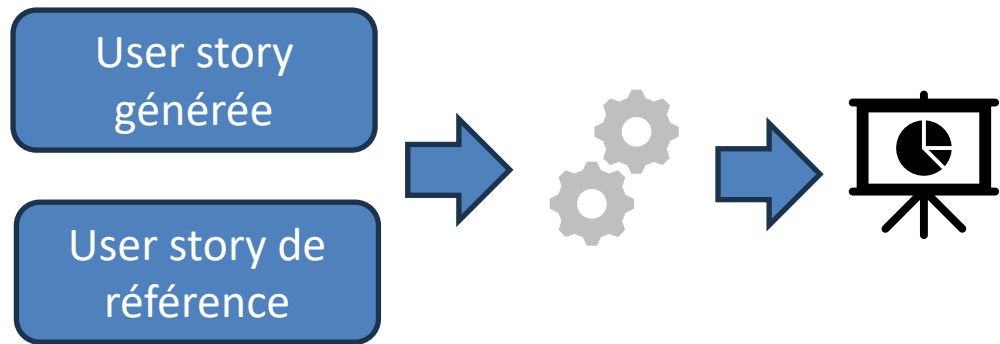
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

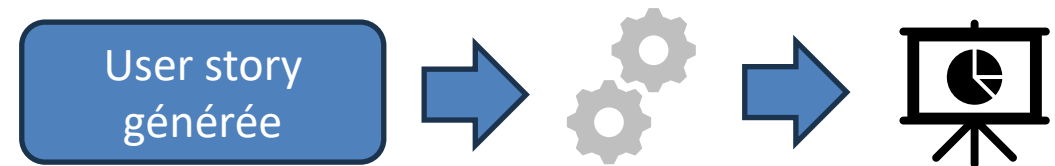
Contrôle Qualité



Supervisées



Non-Supervisées



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Contrôle Qualité



Attributs	Métriques Supervisées	Métriques Non-Supervisées
Fiabilité	Elevée	Contextuelle
Complexité	Faible	Elevée
Coûts additionnels (annotation)	Elevés	Aucuns
Usage pour les industriels	Nul	Contextuel



- Les métriques supervisées sont un standard quantitatif
- Toutefois les industriels n'ont pas la donnée pour travailler en supervisé

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Contrôle Qualité



Métrique	Automatique et Quantitative	Non-supervisée	Prise en compte du contexte projet
INVEST	Red	Green	Green
QUS	Yellow	Green	Yellow
RUST	Red	Green	Red
ISO25010	Red	Green	Red
USQA	Green	Green	Red
ROUGE	Green	Red	Red
CŒUR (ma thèse)	Green	Green	Green



La plupart des métriques disponibles sont insatisfaisantes pour des usages industriels

Autres Métriques NLP possibles :

- BLEU
- BERTScore
- ...

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

Vitesse de pointe

La performance des systèmes IA est un enjeu majeur :

- D'usage
- De productivité
- De fiabilité

Comme toute système critique, ils doivent être testés en charge

un fonctionnement radicalement différent



- **Parcours non déterministes** (conversation, variabilité des prompts)
- **Consommation de ressources variable** (tokens, GPU, contexte)
- **Coûts d'exécution élevés**
- **Temps de réponse variables** (dépendant du contenu)
- **Goulots d'étranglement nouveaux** (GPU, base vectorielle, inférence)
- **Sécurité des données**

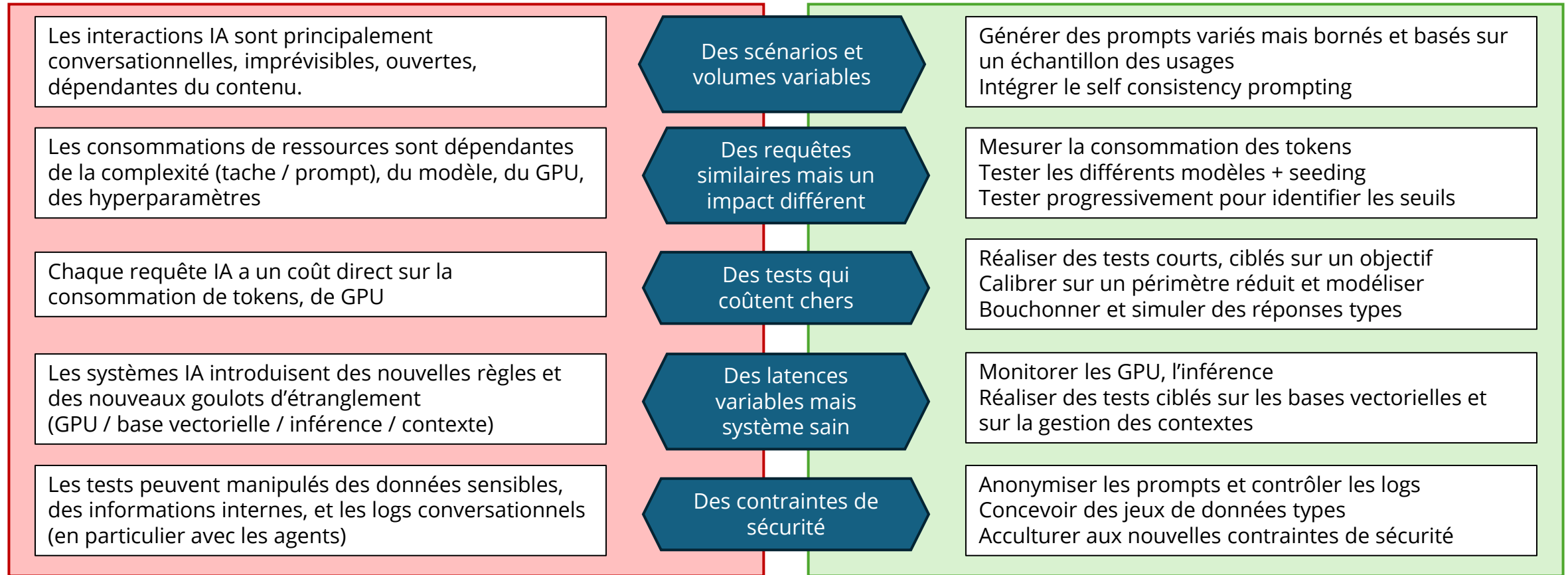


Les systèmes IA ne se testent pas comme les applications d'hier :

→ Ils imposent une stratégie de performance nouvelle, pensée pour des parcours non déterministes, une consommation de ressources instable, des coûts élevés et des points de saturation inédits.

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques



Tester les systèmes IA, c'est : **Exploiter des outils d'observabilité adaptés aux systèmes IA**

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

2 - Maîtriser l'IA par le pilotage et les métriques

La casquette du tableau de bord

OBSERVABILITÉ

Les tests de performances apportent des bénéfices essentiels ...

- Mesurer les performances et valider la robustesse du système
- Identifier les limites de capacité et calibrer le système
- Détecter les goulots d'étranglement spécifique à l'IA
- Estimer les coûts par requête et anticiper le budget de production
- Vérifier la qualité des réponses en charge
- Vérifier la sécurité et la résilience
- Sécuriser le déploiement en production
- Préparer l'intégration dans les outils du quotidien

... Mais les systèmes IA sont vivants, en constante évolution et soumis à un usage imprévisible !



Les tests ne suffisent pas, il faut une observabilité adaptée

Un système IA peut dériver

- **En qualité** → hallucinations, réponses incohérentes
- **En performance** → latence fonction du prompt
- **En coût** → tokens qui explosent
- **En comportement** → Dérive au fil des conversations

- **Performance** : Latence / Temps d'embedding / Temps de recherche vectorielle
- **Coût** : Tokens (type et modèle LLM) / Coût moyen par utilisateur et par usage
- **Gaspillage** : Questions et réponses identiques / Incohérence des prompts ou prompts non adaptés
- **Qualité** : Répétition et abandon / Pertinence et cohérence de la réponse / Hallucinations
- **Sécurité** : Erreurs de sécurité
- **Stabilité** : Taille des réponse / Changement de comportement

Exemples

OBSERVABILITÉ

Observabilité Technique
(Finetuning & Développement d'Agents)



LangSmith



Weights
& Biases



DATADOG

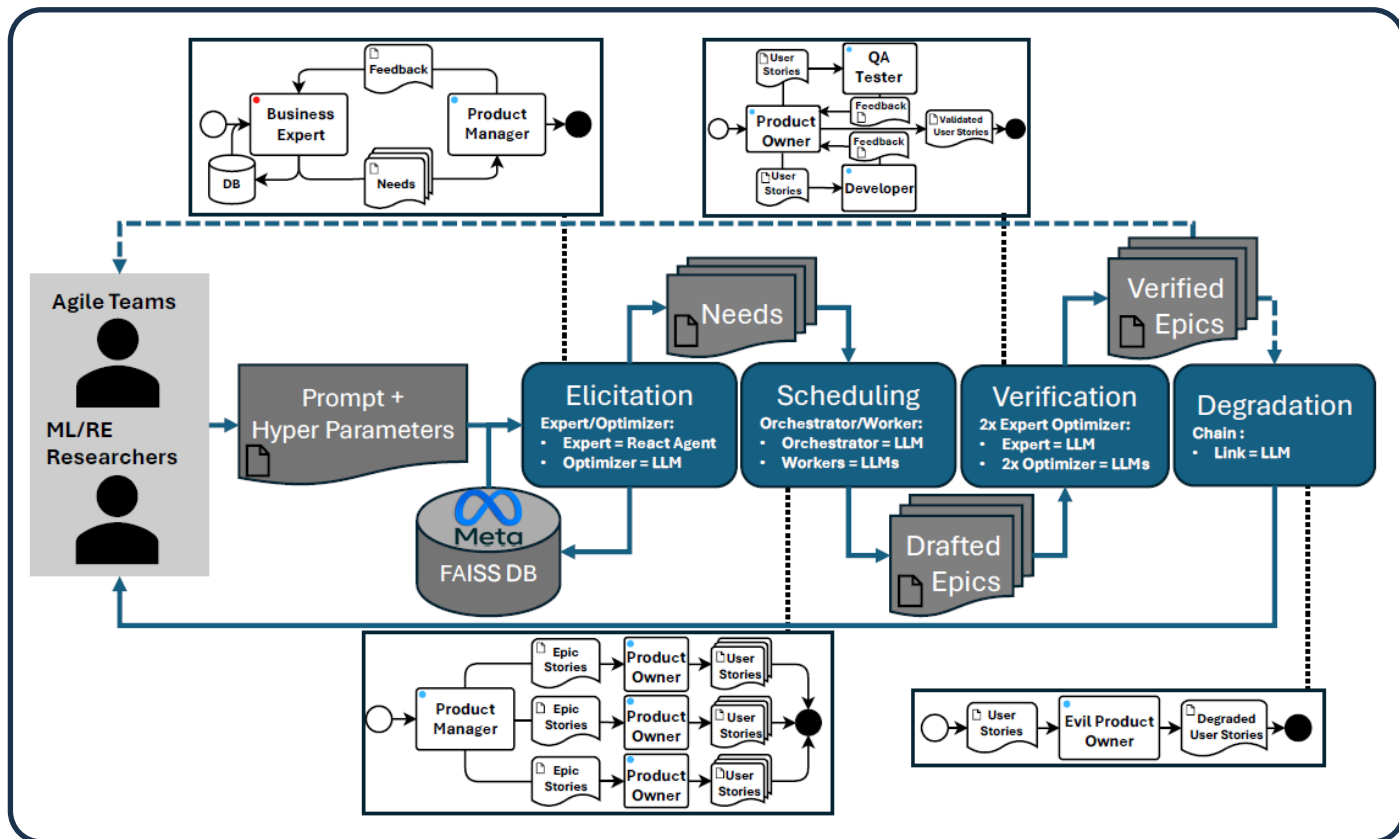


dataiku

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

3 - Retour d'expérience et trajectoires

Text2Stories et CŒUR Perspective recherche et qualité logicielle















✓ Découpage Fonctionnel

✓ Génération de Critères d'Acceptation

✓ Usage de Métriques Standards

✓ Approche Multi-Agent

Evaluation qualitative

Modèles	Découpage Fonctionnel (epics)	Génération de Critères d'Acceptation	Usage de Métriques Standards	Approche Agentique
GeneUS <i>(Rahman et al. 2024)</i>				
UStAI <i>(Yamani et al. 2025)</i>				
Text2Stories <i>(notre modèle)</i>				

Evaluation quantitative



COEUR

**Cohesion & Exhaustiveness
of User-story Representations**



Ensemble de deux métriques
Cohésion et Exhaustivité



Evaluation Quantitative, Non-
supervisée et basée sur le
contexte projet



Assistance à la génération de
Backlog par l'humain et le LLM



**Corrèle à +80% avec des
métriques supervisées
standards**

Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

3 - Retour d'expérience et trajectoires

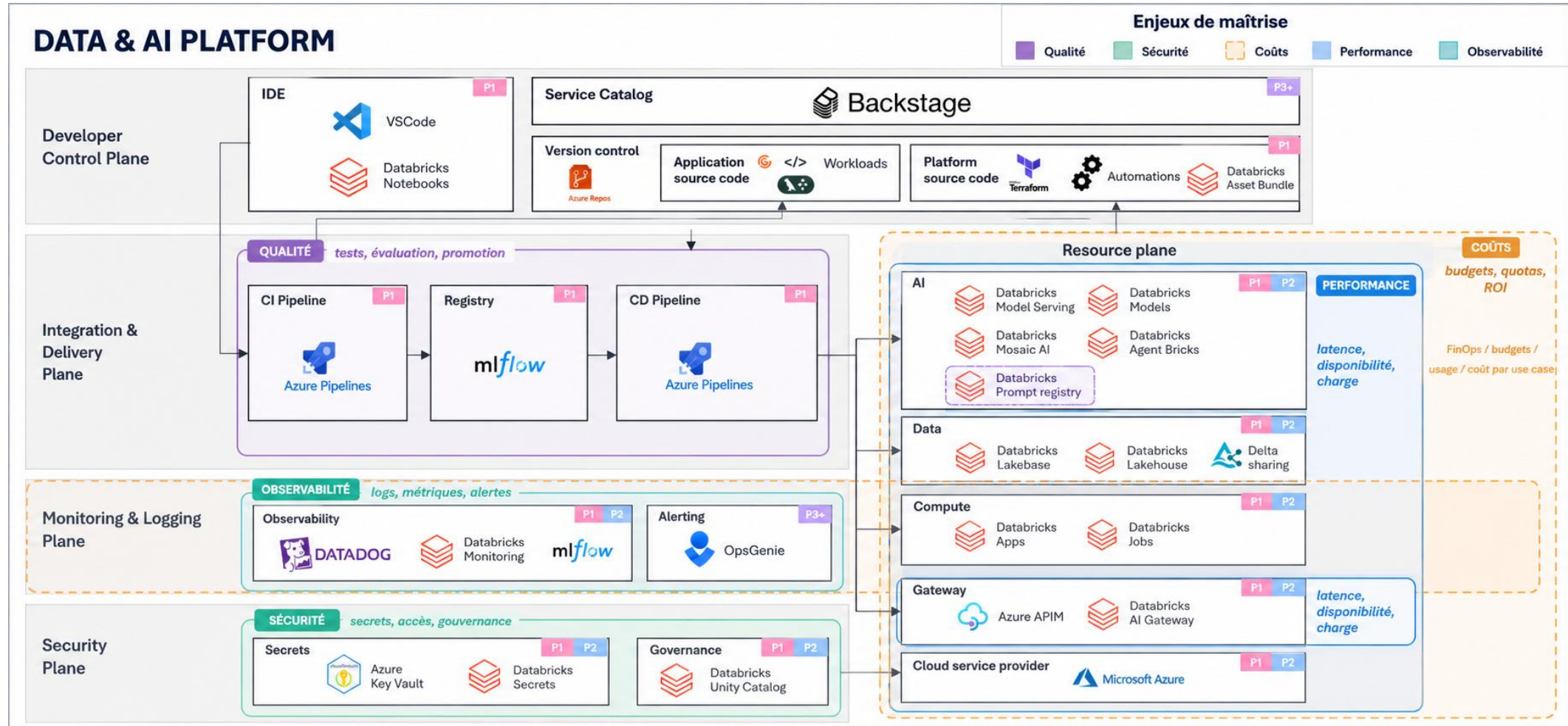
Text2Stories



CŒUR



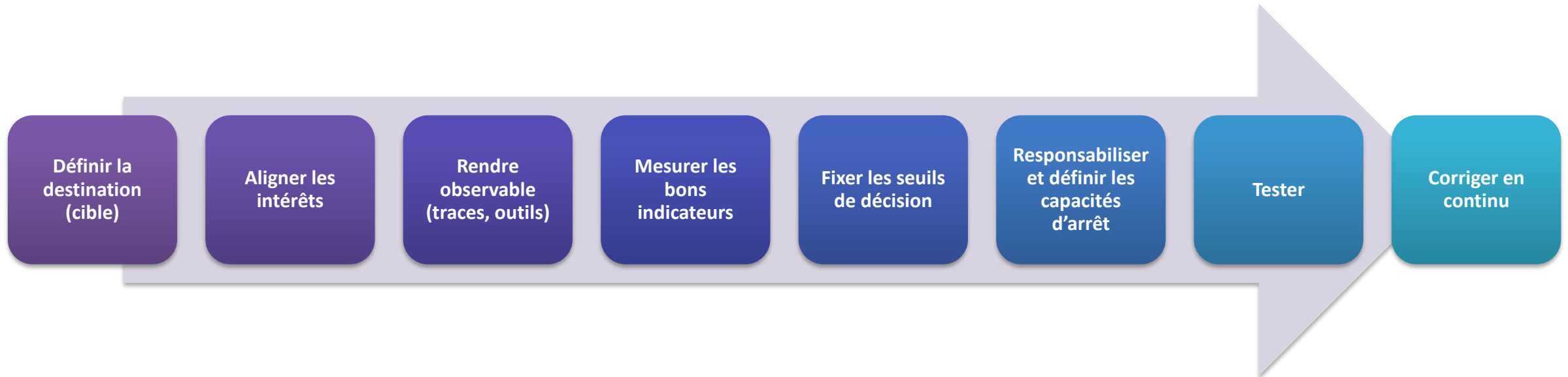
Exemple d'implémentation client outillée azure (en cours de finalisation)



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

Conclusion

Le permis IA ne ralentit pas : il permet d'aller plus loin



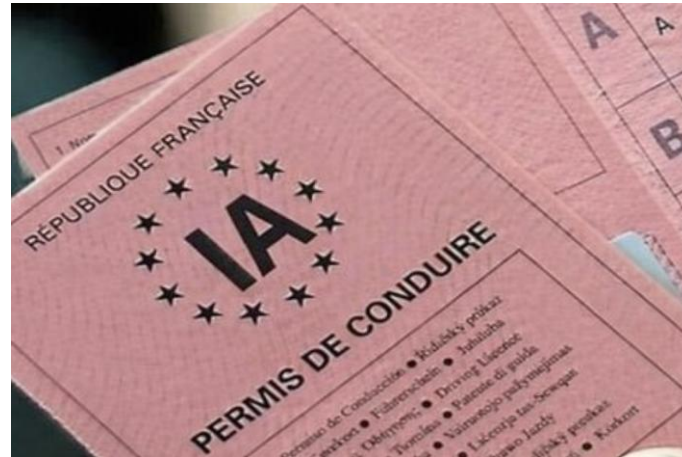
La maîtrise a un coût. Mais l'absence de maîtrise également.
Ce n'est pas simple mais on peut la construire progressivement, usage par usage, en partant d'une intention claire, puis en ajoutant de la visibilité, des règles, des responsabilités, des tests et une boucle d'amélioration continue



Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

Conclusion

Le permis IA ne ralentit pas : il permet d'aller plus loin



Mieux piloter : Voir les **usages**, mesurer les bons **indicateurs**, fixer les **seuils** et **décider**.

Éviter les amendes : Limiter les dérives de **coûts**, de **sécurité**, de **qualité**, de **conformité** et de **confiance**.

Rouler en toute confiance : Construire progressivement des usages IA **arrêtables**, **responsables**, **testés** et **utiles**.

Merci pour votre attention