

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

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



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

Introduction

Le compteur tourne

Les modèles économiques basculent



Requêtes

Tokens

Crédits

Actions
agentiques

Quotas

limites / pay-
as-you-go

Le coût de l'IA ne se limite plus au prix d'un modèle ou d'un abonnement.
Il se déplace dans les crédits, les tokens, les actions réalisées par les agents, les appels d'outils, les traces et les dépassements.

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.

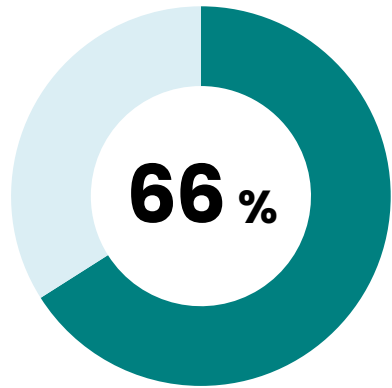
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

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

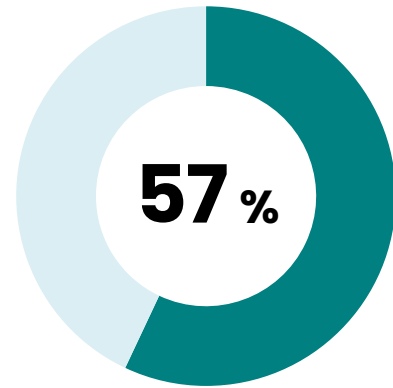
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

1 - L'IA en libre circulation

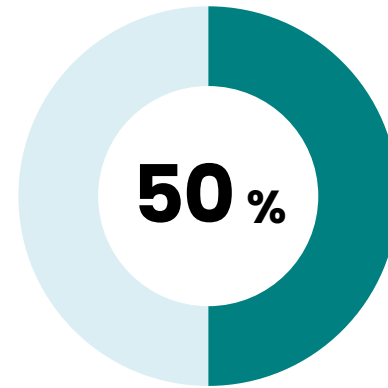
Absence de signalisation ?



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

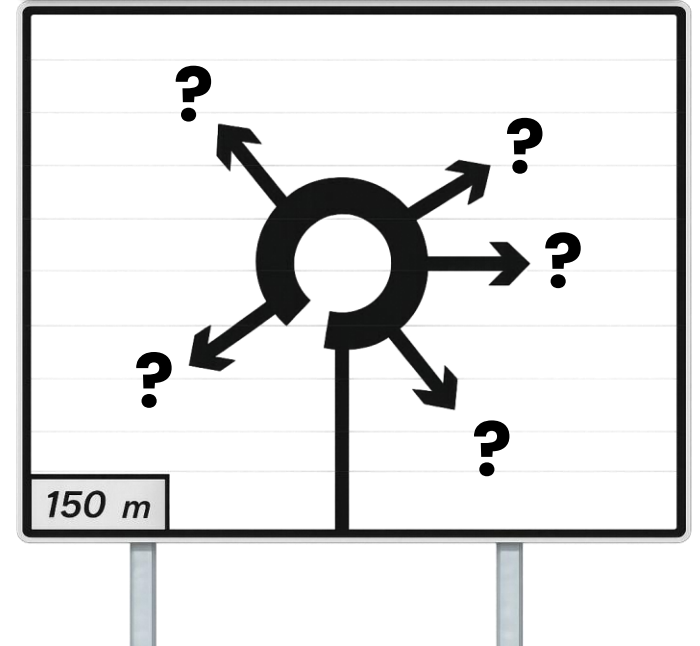


déclarent **cache** leur
usage de l'IA au travail



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

- Usages dispersés
- Comptes personnels
- Outils non validés
- Peu de règles communes
- Peu de formation
- Responsabilités floues



Sources : KPMG – Trust, attitudes and use of AI (2025)

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 DISTINGUER LES USAGES

DÉPASSEMENTS BUDGÉTAIRES DIFFICILES À ATTRIBUER

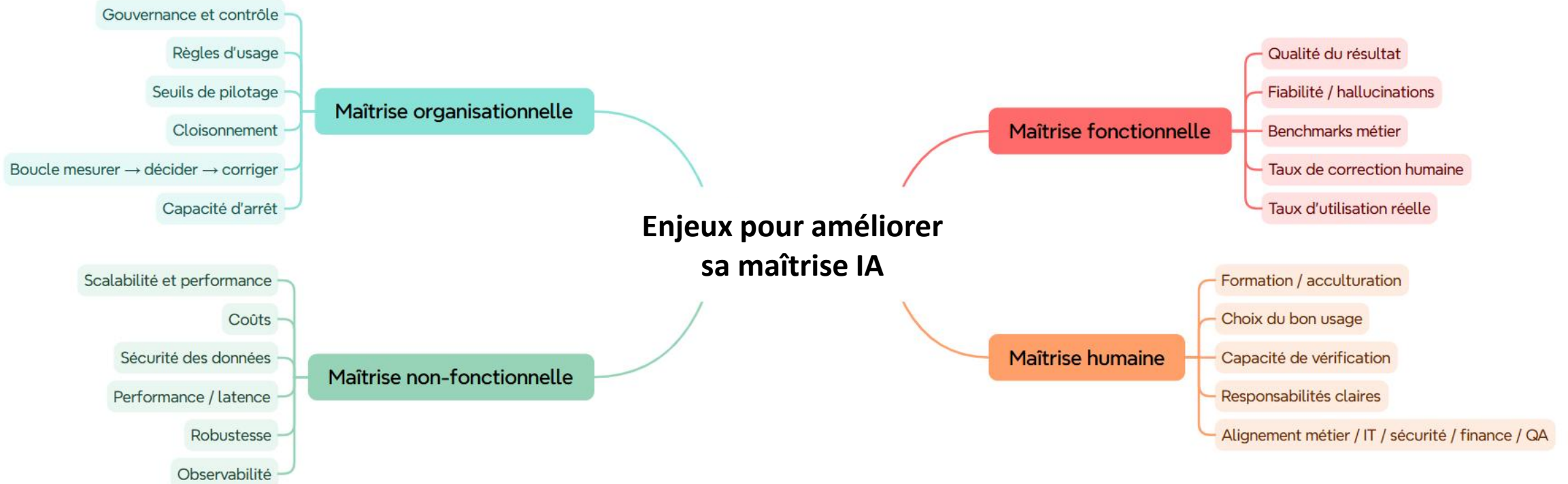
IMPACT GLOBAL EN CAS D'INCIDENT

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

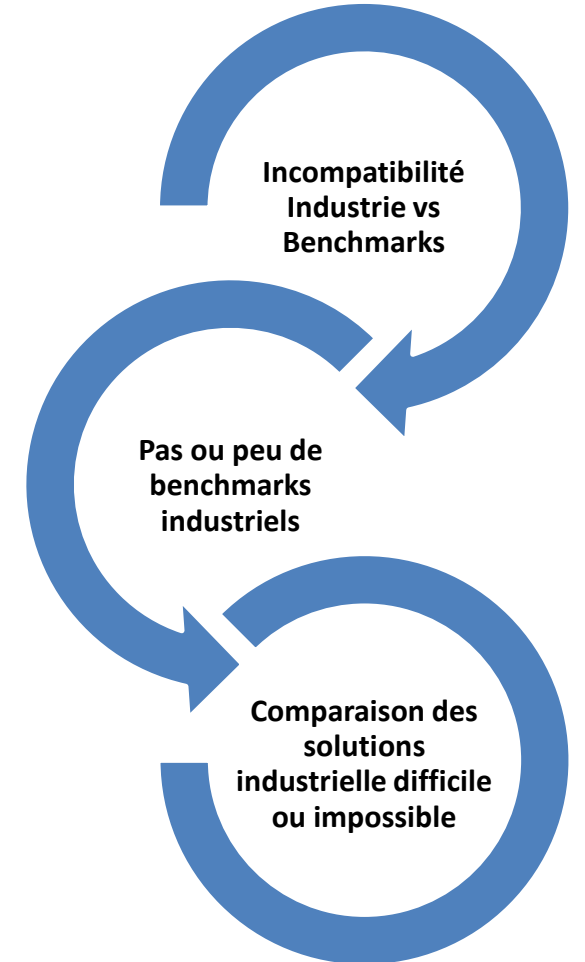
Model	Organization	Global Average	Reasoning Average	Coding Average	Agentic Coding Average
GPT-5 High	OpenAI	79.33	98.17	77.10	46.67
GPT-5 Medium	OpenAI	78.85	96.58	75.05	50.00
GPT-5 Pro	OpenAI	78.73	96.67	72.11	43.33
Claude Sonnet 4.5 Thinking	Anthropic	78.26	95.28	80.36	50.00

LiveBench (White et al. 2024)



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



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



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**

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 surtout celui des résultats et de leur valeur réelle.



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

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

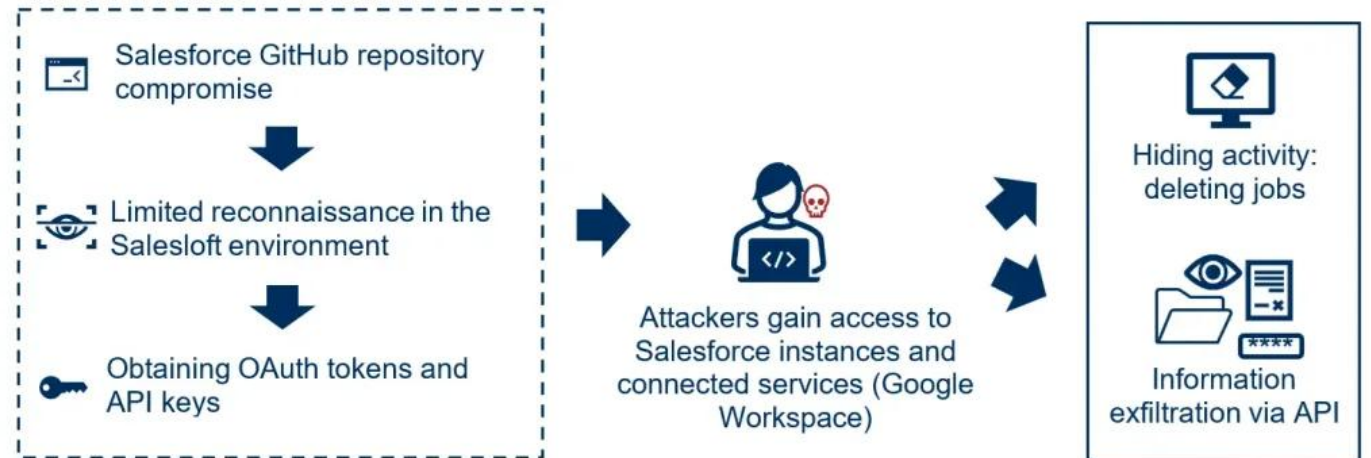
L'importance des glissières

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** : 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

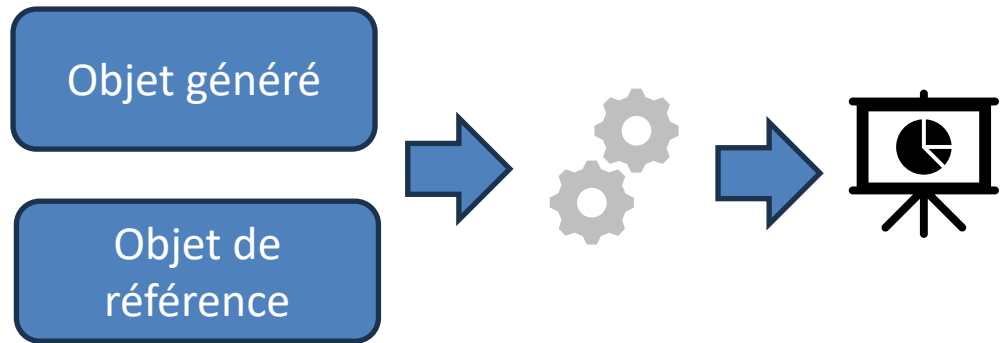
Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

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

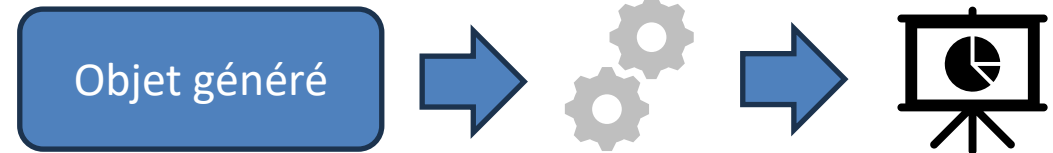
Contrôle Qualité



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 tout système critique, ils doivent être testés en charge

... mais 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**

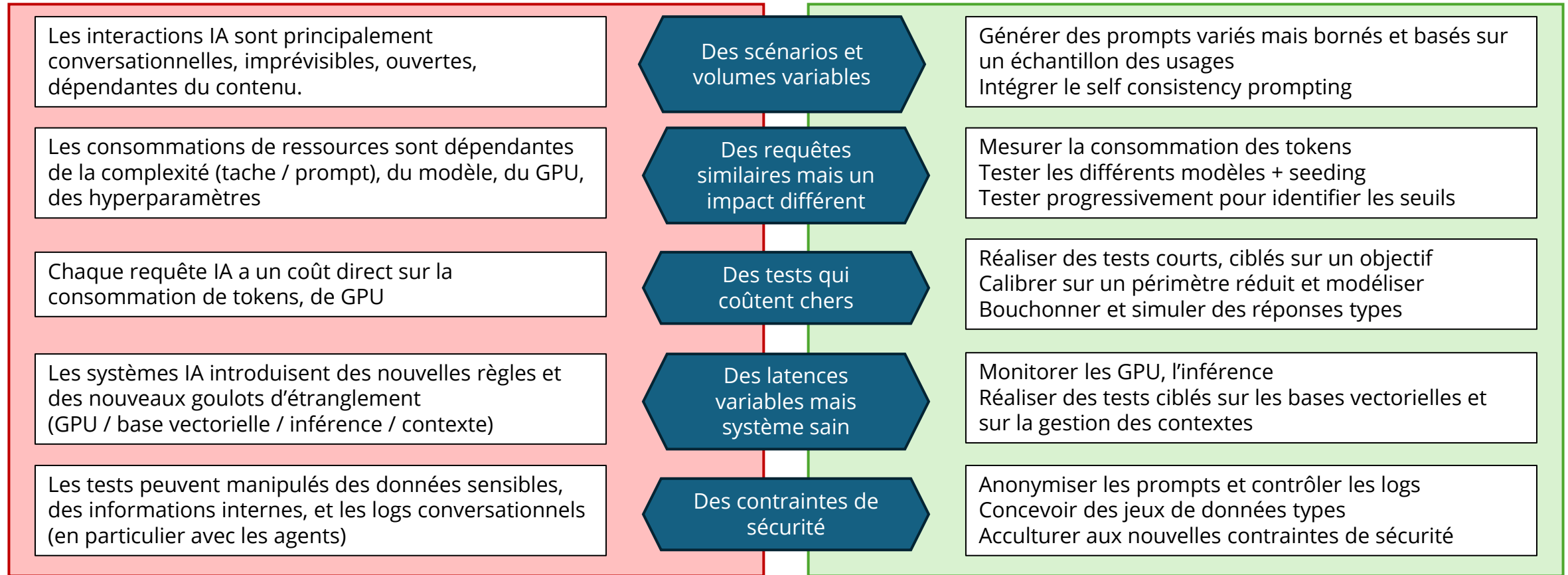


Les systèmes IA ne se testent pas comme les applications d'hier :

→ 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 :

- Utiliser l'IA pour préparer, exécuter, analyser les tests
- 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 cout** → 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

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

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

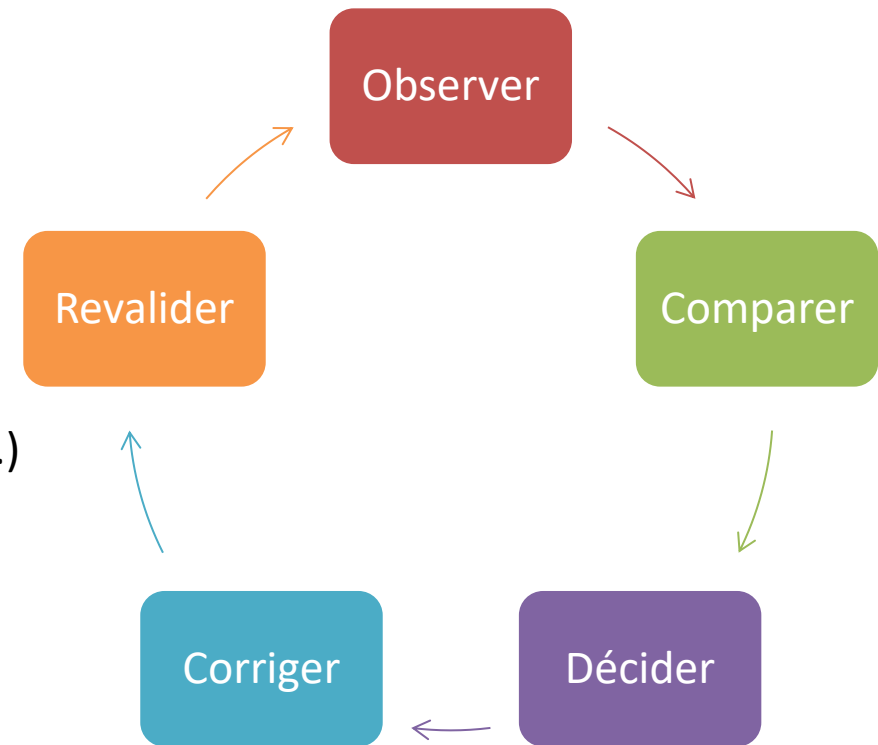
Contrôle technique régulier

OBSERVABILITÉ

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.

1. **Observer** : (usages, gaspillage, coûts, erreurs, latences, données...)
2. **Comparer** : (budget, qualité, performance, autonomie...)
3. **Décider** : (ajuster, limiter, changer, désactiver, industrialiser...)
4. **Corriger** : (prompt, modèle, droits, seuils, validations, actions...)
5. **Revalider** : (benchmark, tests, revue humaine, suivi...)

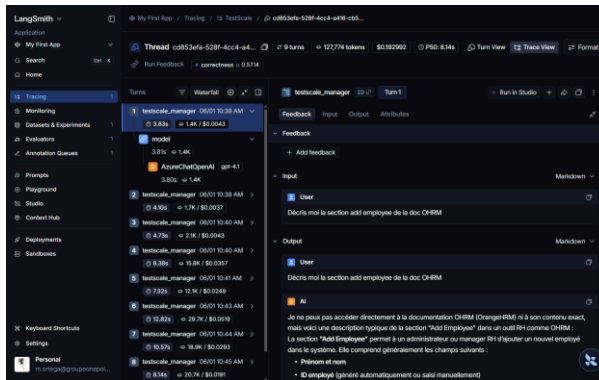
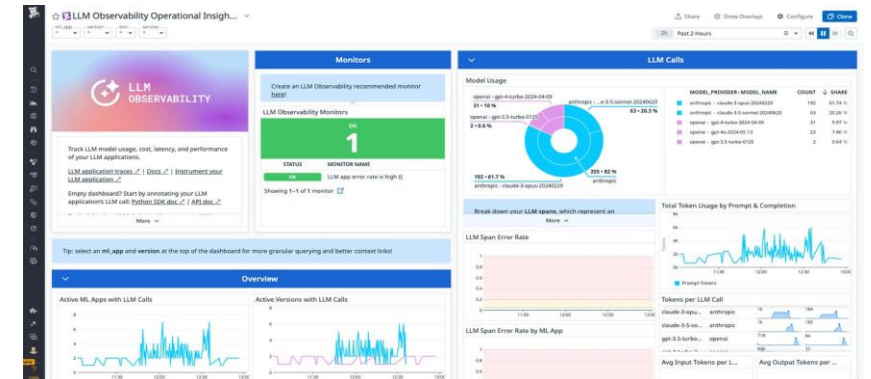
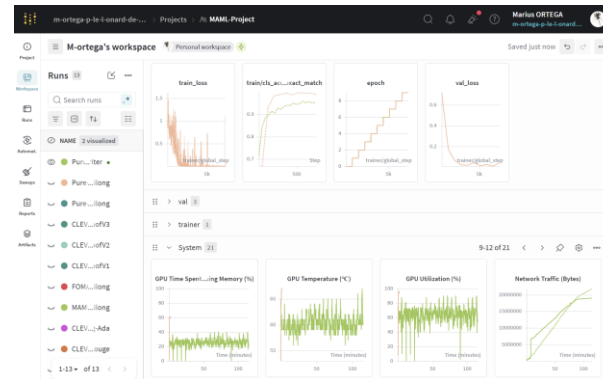
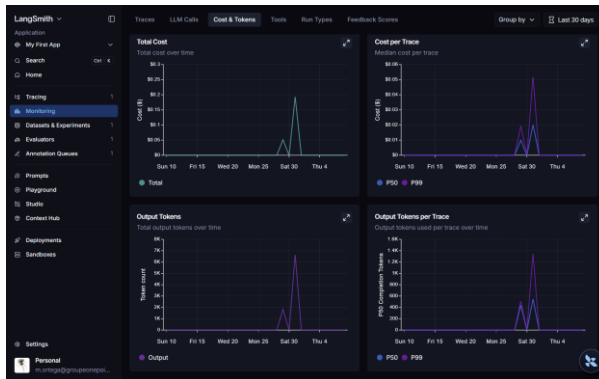


La maîtrise de l'IA n'est pas un contrôle initial. C'est la capacité à voir une dérive, décider quoi faire, et reprendre la main avant qu'elle ne devienne structurelle.

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

3 - Retour d'expérience et trajectoires

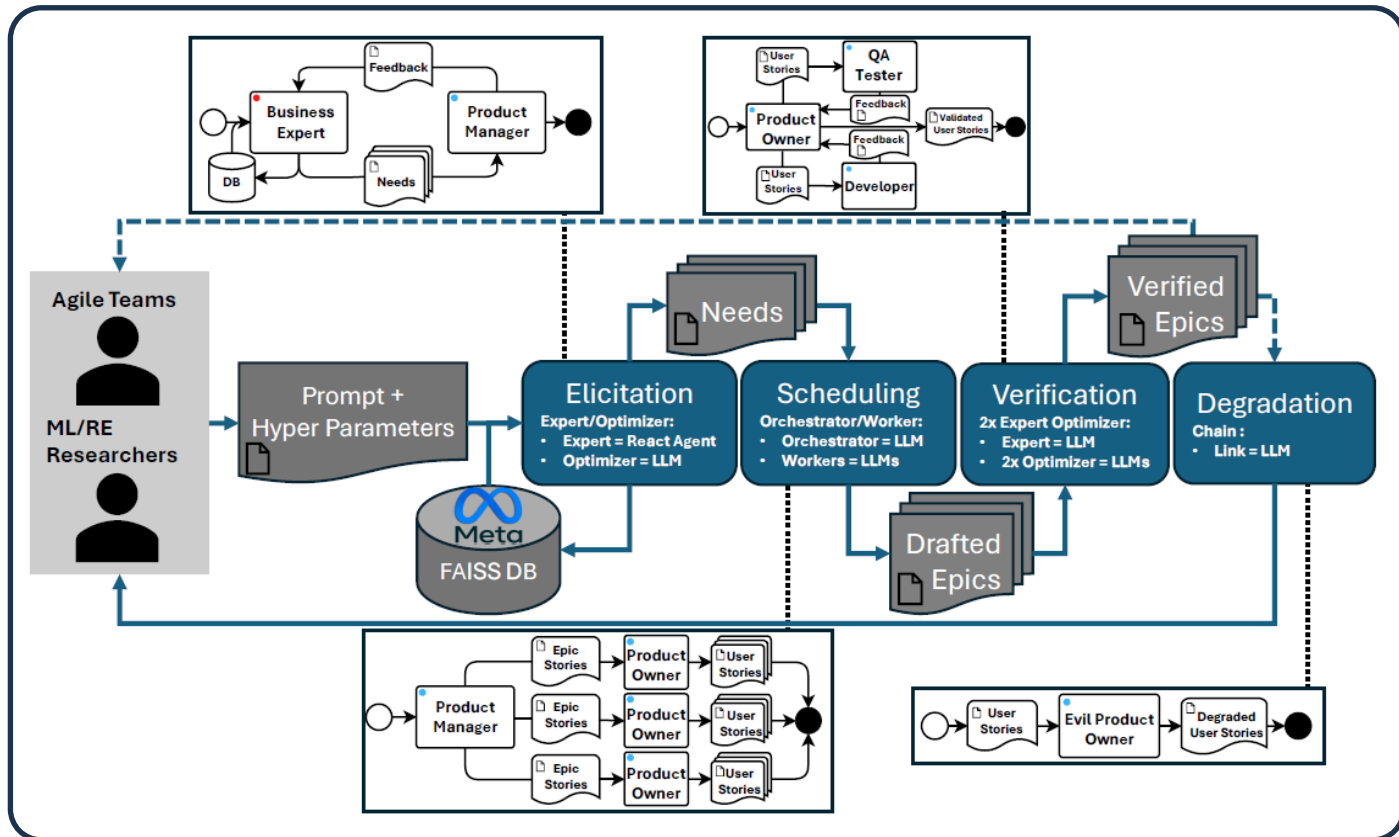
Observabilité et Monitoring : Exemple de Langsmith/Datadog/Weights and Biases



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

3 - Retour d'expérience et trajectoires

Text2Stories pour la Génération et l'Évaluation Automatiques de User Stories



✓ Découpage Fonctionnel

✓ Génération de Critères d'Acceptation













✓ Usage de Métriques Standards

✓ Approche Multi-Agent

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

3 - Retour d'expérience et trajectoires

Evaluation qualitative de Text2Stories

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>				

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

3 - Retour d'expérience et trajectoires

Evaluation Quantitative de Text2Stories

Méthode	Alignement Humain-Modèle				Défauts de Format et d'Unicité				
	INVEST (Wake, 2003)				AQUSA Score (QUS) (Lucassen et al., 2016)				
	Acc ↑	Prec ↑	Recall ↑	F1 ↑	Atom. ↓	Well-Fo. ↓	Mini. ↓	Unif. ↓	Uniq. ↓
LLM	0.93	0.93	1.00	0.96	0.00	0.00	0.00	0.00	0.54
LLM + CoT	0.73	0.75	0.93	0.82	0.20	0.00	0.04	0.12	0.00
RAG + CoT	0.77	0.82	0.89	0.85	0.16	0.00	0.00	0.12	0.00
T2S (Gen)	0.83	0.85	0.96	0.90	0.08	0.00	0.00	0.42	0.12

400+ user stories
20+ consultants

- ?
- ✓
- ✓
- ✓



Pourquoi le LLM simple obtient-il un meilleur score que les modèles plus avancés

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

3 - Retour d'expérience et trajectoires

Les Grandes Problématiques en Requirement Engineering : « Sur mon PC, ça fonctionne »

1 Précarité des Benchmarks :
« Je suis meilleur que toi, mais pas sur les mêmes données »

Home > Requirements Engineering > Article **383 cites**

Improving agile requirements: the Quality User Story framework and tool

RE 2015 | Open access | Published: 01 April 2016
Volume 21, pages 383–403, (2016) [Cite this article](#)

Garm Lucassen, Fabiano Dalpiaz, Jan Martijn E. M. van der Werf & Sjaak Brinkkemper

Jeu de Données :

- 1023 User Stories
- 18 Projets
- Privé



arXiv > cs > arXiv:2306.12132 **50 cites**

Computer Science > Software Engineering

[Submitted on 21 Jun 2023]

ChatGPT as a tool for User Story Quality Evaluation: Trustworthy Out of the Box?

Krishna Ronanki, Beatriz Cabrero-Daniel, Christian Berger

Jeu de Données :

- 3580 User Stories
- 8 Projets
- Public

3 Tâches d'apprentissage mal définies :
« On fait la même chose, mais différemment »

Fonctionnalité	GeneUS (Rahman et al., 2024)	UStAI (Yamani et al., 2025)	T2S (La Nôtre)
Génération Contextualisée (US)	×	×	✓
Génération d'AC	✓	×	✓
Évaluation via Métriques Std.	×	✓	✓

2 Désalignement des métriques :
« On est bons, mais pas sur la bonne tâche »

Metric	Required for COEUR Comparison		Not Required Project Context Aware
	Automatic and Quantitative	Unsupervised	
INVEST [35]	×	✓	✓
QUS [20, 21]	○	✓	○
RUST [27]	×	✓	×
ISO25010 [10, 11]	×	✓	×
USQA [17]	✓	✓	×
Santos [8]	✓	×	×
COEUR (Ours)	✓	✓	✓

✓: Yes, ○: Partially, ×: No

Conferences > 2013 2nd SEMAT Workshop on a ...

Speeding-up software engineering's escape from its pre-paradigmatic stage

Publisher: IEEE

[Cite This](#) [PDF](#)

laakov Exman All Authors

3 Cites in Papers | 90 Full Text Views

[R](#) [↩](#) [©](#) [📁](#) [🔔](#)

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

3 - Retour d'expérience et trajectoires

CŒUR : Une mesure non-supervisée de la qualité des Backlogs produits



Accepté pour publication

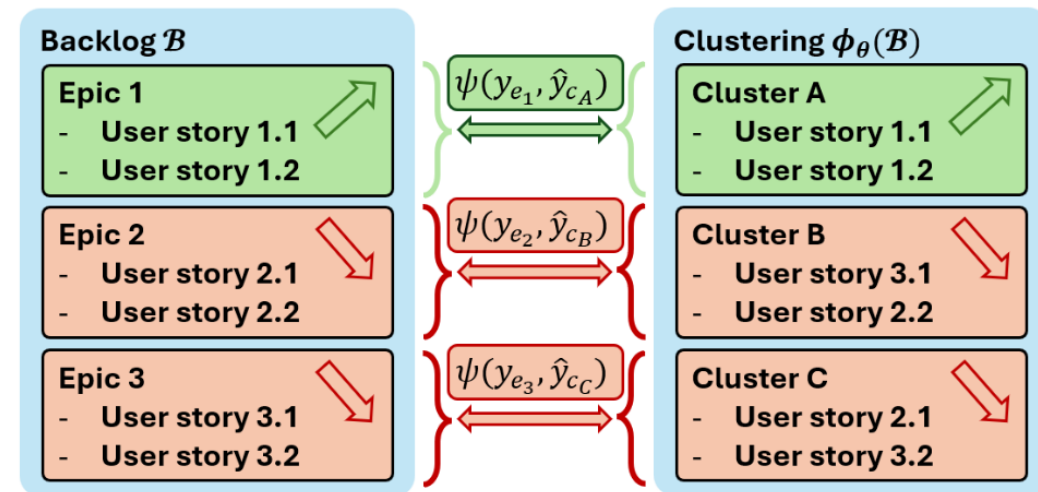
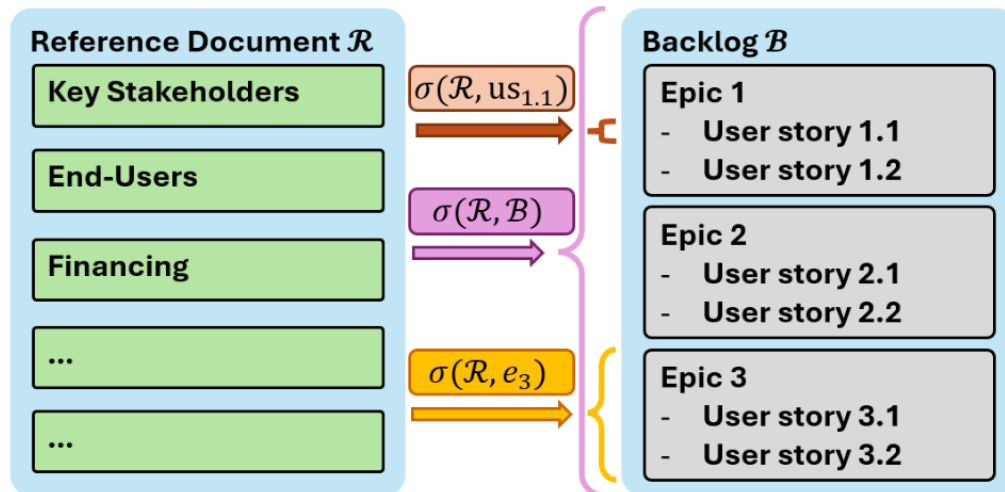
Objectif : Rendre compte des 2 principales activités d'un Product Owner ?

L'exhaustivité :

Le backlog retranscrit-il fidèlement les spécifications d'un point de vue sémantique ?

La cohésion :

Le backlog est-il correctement maintenu ? Les fonctionnalités (epics) sont-elles disjointes ? Les user stories sont-elles liées à l'epic la plus appropriée sémantiquement ?



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

3 - Retour d'expérience et trajectoires

Evaluation Quantitative de CŒUR

Meilleur que les métriques non-supervisée spécifiques à ce problème de Software Engineering

		ROUGE Metrics			BERTScore Metrics		
		R-1	R-2	R-L	B-P	B-R	B-F1
GPT-4.1	COEUR	0.846***	0.790**	0.748**	0.503	0.664*	0.650*
	Coh	0.622*	0.573	0.566	0.182	0.336	0.350
	Exh	0.860***	0.769**	0.762**	0.587*	0.804**	0.769**
Mistral-22b	COEUR	0.839***	0.804**	0.818**	0.832***	0.839***	0.867***
	Coh	0.643*	0.692*	0.650*	0.713**	0.462	0.636*
	Exh	0.769**	0.727**	0.748**	0.727**	0.748**	0.762**
GPT-OSS	COEUR	0.867***	0.818**	0.804**	0.629*	0.734**	0.727**
	Coh	0.629*	0.594*	0.685*	0.601*	0.545	0.559
	Exh	0.699*	0.671*	0.559	0.448	0.678*	0.650*

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; R-1/2/L: ROUGE-1/2/L;
B-P/R/F1: BERTScore Precision/Recall/F1

	External Noise Experiment (Section 4.2.1)					Internal Noise Experiment (Section 4.2.2)					
	QUS [21]	USQA [17]	Exh (Ours)	Coh (Ours)	COEUR (Ours)	QUS [21]	USQA [17]	Exh (Ours)	Coh (Ours)	COEUR (Ours)	
Trident	$\rho \uparrow$	1.00±0.00	0.97±0.02	1.00±0.00	0.94±0.02	0.96±0.01	-0.00±0.17	0.00±0.00	0.00±0.00	0.95±0.03	0.95±0.03
	$r_s \uparrow$	1.00±0.00	0.98±0.01	1.00±0.00	0.94±0.02	0.98±0.01	0.07±0.34	0.00±0.00	0.00±0.00	0.94±0.05	0.94±0.05
	$p(\rho) \downarrow$	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.59±0.29	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
	$p(r_s) \downarrow$	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.25±0.25	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Retro	$\rho \uparrow$	-0.96±0.02	0.65±0.18	1.00±0.00	0.88±0.07	1.00±0.00	0.04±0.06	0.00±0.00	0.00±0.00	0.98±0.01	0.98±0.01
	$r_s \uparrow$	-0.96±0.03	0.62±0.18	1.00±0.00	0.88±0.08	1.00±0.00	0.15±0.17	0.00±0.00	0.00±0.00	0.97±0.01	0.97±0.01
	$p(\rho) \downarrow$	0.00±0.00	0.01±0.03	0.00±0.00	0.00±0.00	0.00±0.00	0.52±0.39	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
	$p(r_s) \downarrow$	0.00±0.00	0.02±0.03	0.00±0.00	0.00±0.00	0.00±0.00	0.23±0.28	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Alfred	$\rho \uparrow$	0.81±0.12	0.63±0.18	1.00±0.00	0.93±0.02	0.99±0.00	-0.37±0.20	0.00±0.00	0.00±0.00	0.91±0.03	0.91±0.03
	$r_s \uparrow$	0.78±0.16	0.58±0.22	1.00±0.00	0.94±0.02	0.99±0.00	-0.50±0.22	0.00±0.00	0.00±0.00	0.89±0.05	0.89±0.05
	$p(\rho) \downarrow$	0.00±0.00	0.02±0.04	0.00±0.00	0.00±0.00	0.00±0.00	0.16±0.24	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
	$p(r_s) \downarrow$	0.00±0.01	0.06±0.14	0.00±0.00	0.00±0.00	0.00±0.00	0.09±0.23	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00

ρ : Pearson Coeff., r_s : Spearman Coeff., $p()$: p-value, \uparrow : Higher is Better, \downarrow : Lower is Better

Corrélation en monotonie (Spearman) à ~80% avec les baselines supervisées standards.

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

3 - Retour d'expérience et trajectoires

CŒUR : Une mesure non-supervisée de la qualité des Backlogs produits



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



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

Ce qu'il faut éviter :

Généraliser un POC sans benchmark métier représentatif.

Mesurer les coûts sans les rattacher aux usages et à la valeur produite

Confondre sortie convaincante et rapidité avec sortie fiable et productivité

Lancer un agent sans observer les logs.

Centraliser les usages sans distinguer coûts, risques et valeur

Connecter une IA au SI sans cartographier les données accessibles

Ne pas prévoir de capacité d'arrêt.

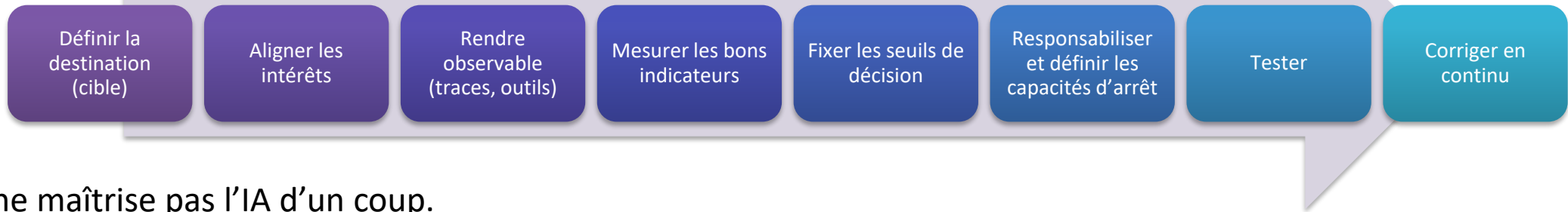


Permis de conduire IA : mieux piloter, éviter les amendes et rouler en toute confiance.

Conclusion

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

La maîtrise de l'IA dans une organisation n'est pas simple et ne s'obtient pas du jour au lendemain. Elle se construit 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



On ne maîtrise pas l'IA d'un coup.

On maîtrise des usages IA, progressivement, jusqu'à pouvoir les piloter à l'échelle.

La maîtrise aussi a un coût. Mais l'absence de maîtrise également. La question à se poser :
À partir de quel volume d'usage l'absence de gouvernance coûte plus cher que sa mise en place ?

Merci pour votre attention