

CEDRIC

LEFEBVRE

Applicabilité de l'IA générative
pour le test des systèmes
embarqués

(focus sur l'aide à l'analyse des exigences)



(expleo)

Think bold, act reliable

JOURNÉE
THÉMATIQUE
IA GÉNÉRATIVE
POUR TESTER



AIDEST

Artificial Intelligence Driven Embedded System Testing



CFTL
Centre Français
des Tests Logiciels

Sommaire

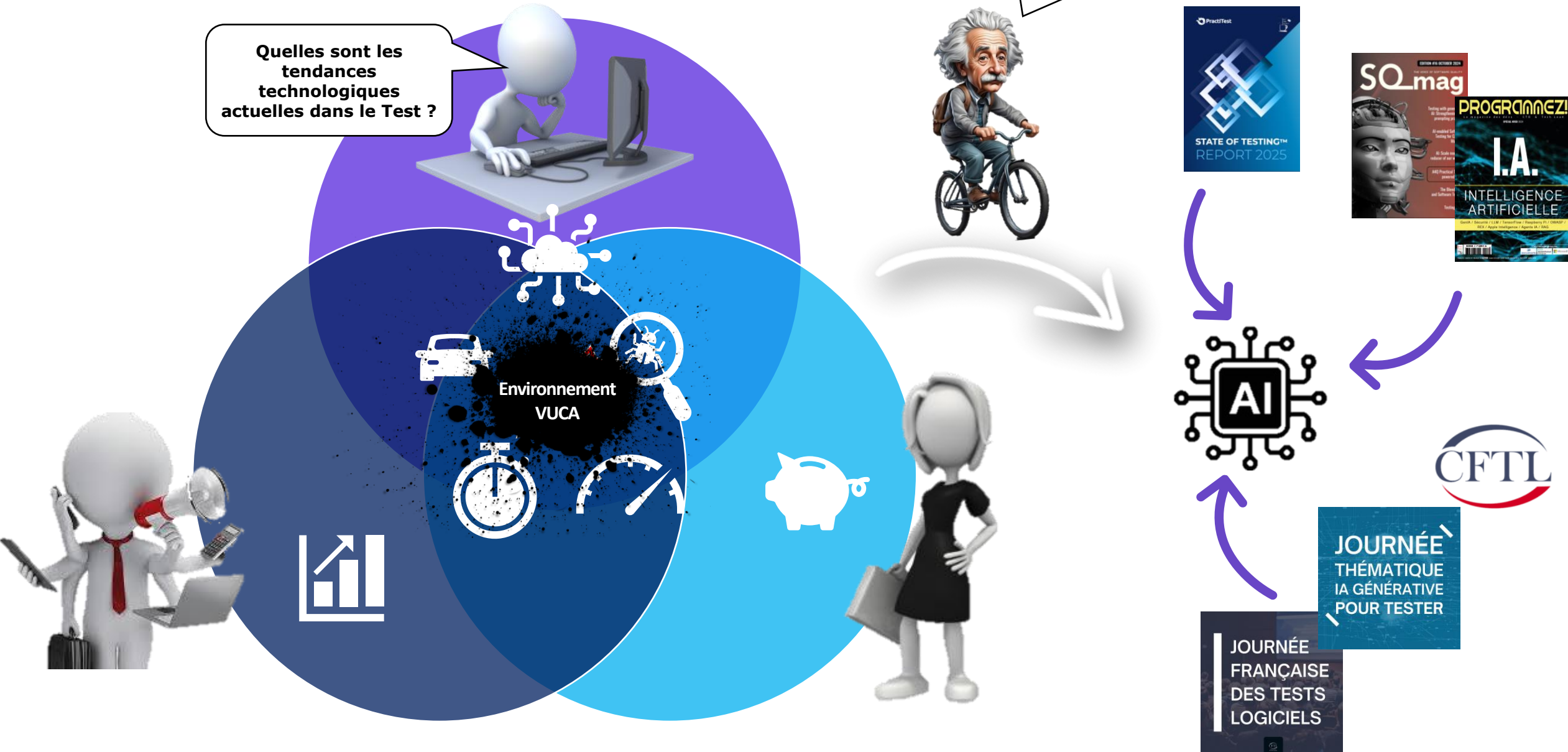
1. Le Pourquoi ou les enjeux du sujet
2. Le Comment ou la méthodologie appliquée
3. Le Quoi ou les résultats obtenus
4. L'heure du bilan et des interrogations

1. Le pourquoi ou les enjeux du sujet

Joindre l'outil à l'agréable.

Quelles sont les
tendances
technologiques
actuelles dans le Test ?

La vie, c'est comme une
bicyclette, il faut avancer pour
ne pas perdre l'équilibre !



1. Le pourquoi ou les enjeux du sujet

Optimiser sous les contraintes.



Artificial Intelligence Driven Embedded System Testing

2. Le Comment ou la méthodologie appliquée

Des contraintes naissent les solutions.



Disponibilité



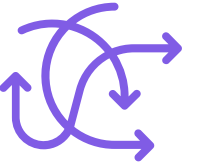
**Manque de compétences
en IA/codage**



**Confidentialité et
sécurité des données**



**Coût élevé de
certaines solutions**



Complexité du sujet



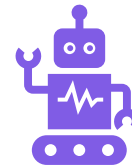
Mise en pratique



**Confidentialité et
sécurité des données**



Autonomie technique



Prioriser les expérimentations.
Influencer le choix des solutions à base d'IA.

2. Le Comment ou la méthodologie appliquée

IA qu'à demander.

Prompt basé sur la structure **Rôle-Contexte-Instructions-Données-Contraintes-Format**.

Tu es mon assistant chef de projet pour m'aider à choisir la meilleure solution technique IA.

Le projet concerne l'étude de l'applicabilité des solutions IA dans le domaine des tests de logiciels et de systèmes embarqués.

Pour chacune de ces solutions techniques, je te demande de :

- Estimer les avantages.
- Estimer les inconvénients.
- Estimer les investissements matériels.
- Estimer les investissements financiers y compris le coût en euros.
- Estimer le temps nécessaire.
- Estimer l'applicabilité au domaine des tests de logiciels et de systèmes embarqués.
- Estimer la maturité

Voici les solutions techniques à analyser :

- Agent IA.
- IA générative par chatbot.
- IA générative par chatbot avec pièce jointe.
- IA générative par chatbot avec fine tuning.
- IA générative par chatbot avec RAG.
- IA générative par API.
- IA générative par API avec pièce jointe.
- IA générative par API avec fine tuning.
- IA générative par API avec RAG

Les contraintes imposées sont :

- Les solutions utiliseront des modèles pré-entraînés.
- Les solutions à base de Chatbot utiliseront la solution existante Copilot Web.

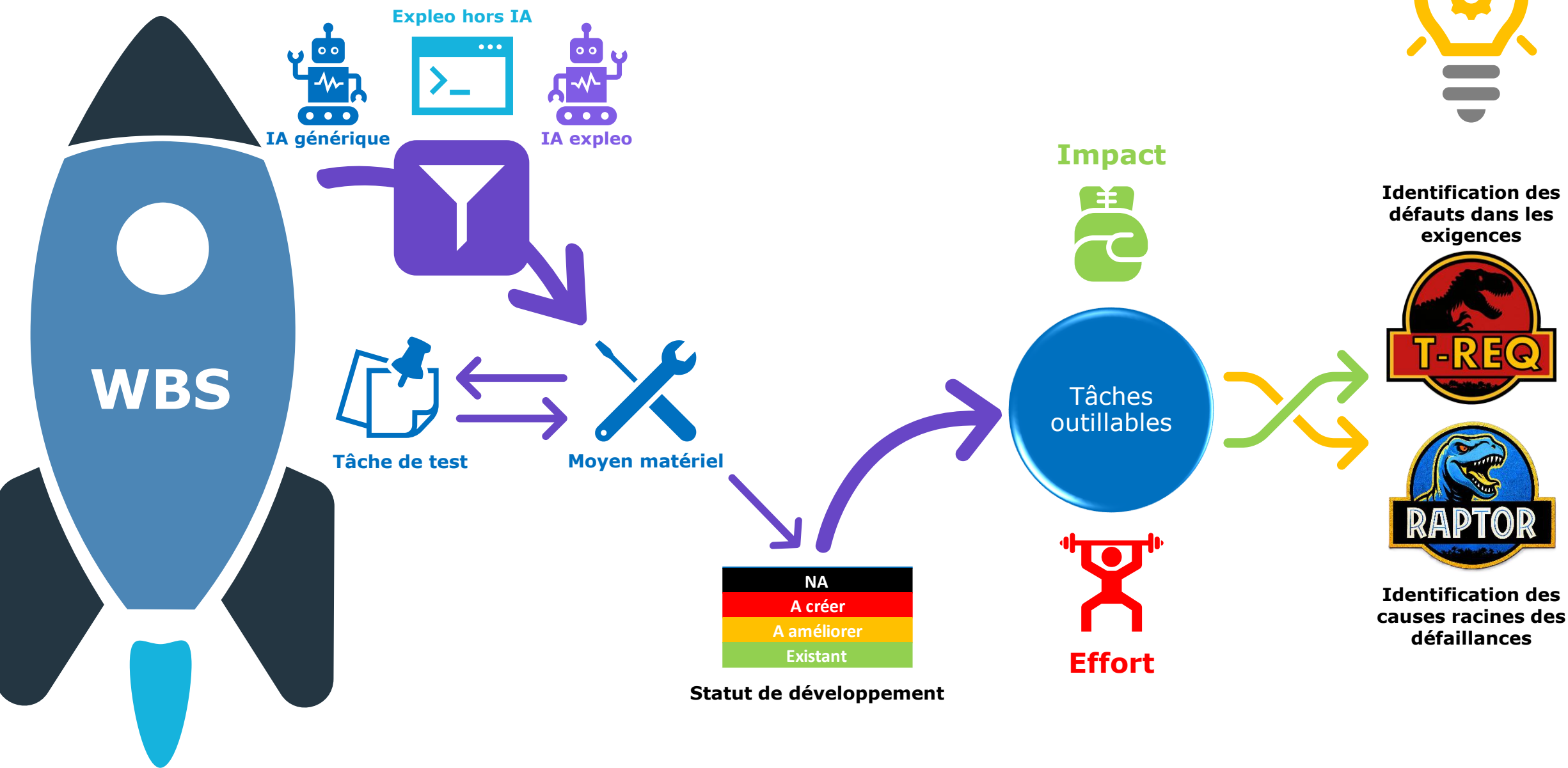
Les résultats seront mis sous forme d'un tableau comparatif.

Solution Technique		Option	Avantages	Inconvénients	Investissements Matériels	Investissements Financiers		Temps Nécessaire	Applicabilité au Test Logiciel		Maturité	Score
IA générative	Agent IA (modèle pré-entraîné)	-	Automatisation des tâches répétitives Amélioration de l'efficacité opérationnelle Interaction en temps réel	Limité par les scénarios prédéfinis Mins flexible face aux requêtes complexes	Serveurs Infrastructure réseau	Développement Intégration	10 000 - 30 000 €	1 à 2 mois	Modérée	Utile pour automatiser des tests répétitifs	Élevée	60%
						API	1 000 - 5 000 € par an					
	chatbot (Copilot web)	-	Réponses personnalisées et contextuelles Amélioration de l'expérience utilisateur	Risque de réponses inappropriées Besoin de supervision humaine	Serveurs	API Copilot	1 000 - 5 000 € par an	Prêt à l'emploi	Modérée	Utile pour des tests interactifs	Modérée	83%
		Pièce jointe	Capacité à traiter et analyser des documents Réponses plus complètes	Complexité accrue Risque de sécurité des données	Serveurs	API Copilot	1 000 - 5 000 € par an	Prêt à l'emploi	Modérée	Utile pour tester des scénarios impliquant des documents	Modérée	83%
		Fine Tuning	Réponses hautement personnalisées Amélioration de la précision des réponses	Coût élevé de l'entraînement Besoin de données spécifiques	Serveurs puissants pour l'entraînement	Entraînement Données	30 000 - 70 000 €	2 à 4 mois	Élevée	Utile pour des tests spécifiques et personnalisés	Modérée	53%
		RAG	Réponses basées sur des données à jour Flexibilité et précision accrues	Complexité de mise en œuvre Besoin de gestion de la base de connaissances	Serveurs pour héberger la base de données vectorielle	API Copilot Gestion des données	3 000 - 12 000 € par an	2 à 4 mois	Élevée	Utile pour tester des scénarios basés sur des données à jour	Modérée	73%
	API (modèles pré-entraînés)	-	Intégration facile avec d'autres systèmes Flexibilité dans l'utilisation des modèles	Dépendance aux fournisseurs d'API Coût récurrent des appels API	Infrastructure réseau	Développement Intégration	10 000 - 30 000 €	1 à 2 mois	Modérée	Utile pour intégrer des tests dans différents systèmes	Élevée	60%
						API	1 000 - 5 000 € par an					
		Fine Tuning	Réponses hautement personnalisées Amélioration de la précision des réponses	Coût élevé de l'entraînement Besoin de données spécifiques	Serveurs puissants pour l'entraînement	Entraînement Données	30 000 - 70 000 €	2 à 4 mois	Élevée	Utile pour tester des scénarios basés sur des données à jour	Modérée	53%
		RAG	Réponses basées sur des données à jour Flexibilité et précision accrues	Complexité de mise en œuvre Besoin de gestion de la base de connaissances	Serveurs pour héberger la base de données vectorielle	API Copilot Gestion des données	3 000 - 12 000 € par an	2 à 4 mois	Élevée	Utile pour tester des scénarios basés sur des données à jour	Modérée	73%



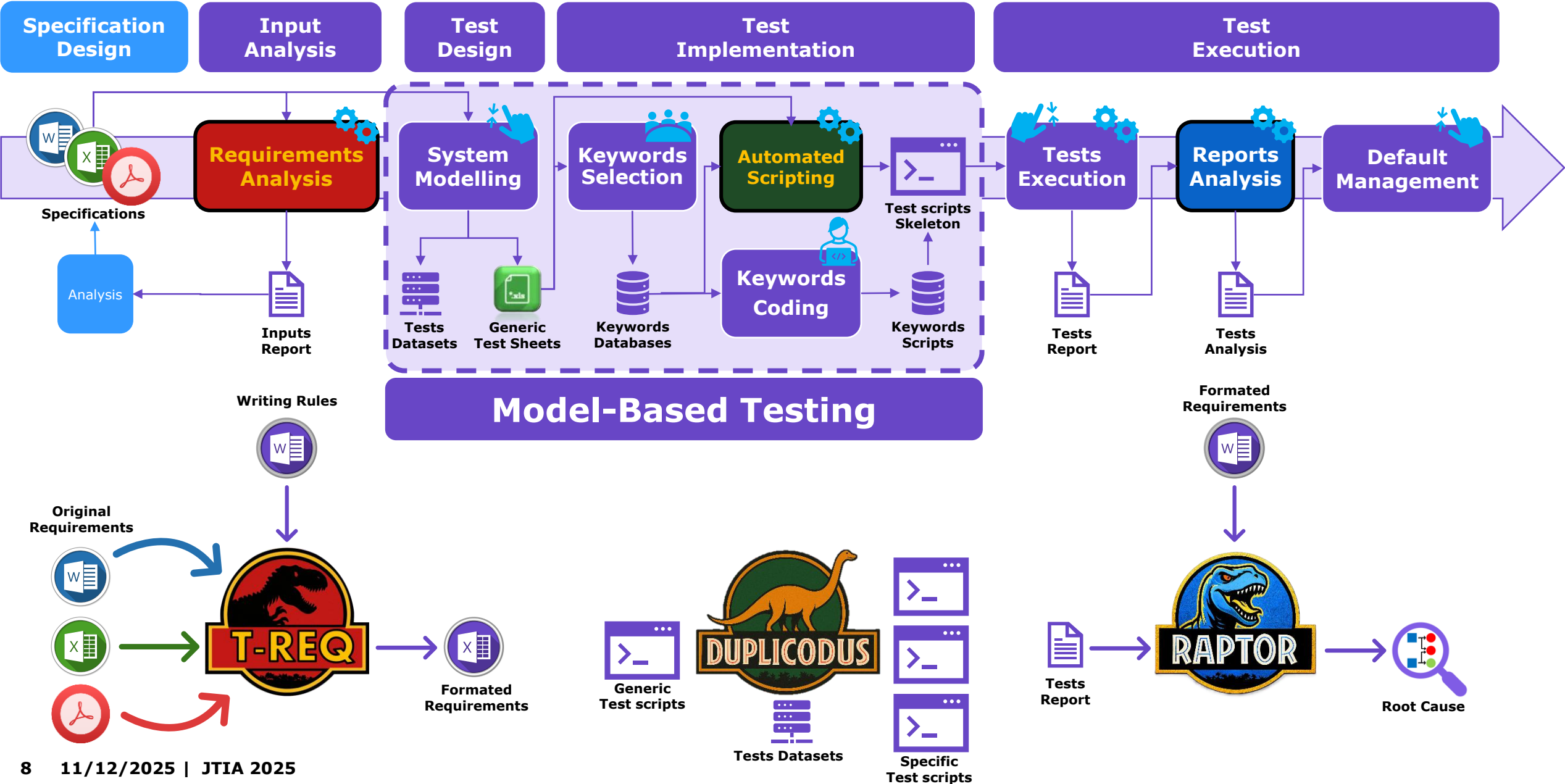
2. Le Comment ou la méthodologie appliquée

Le Jurassic Pack.



2. Le Comment ou la méthodologie appliquée

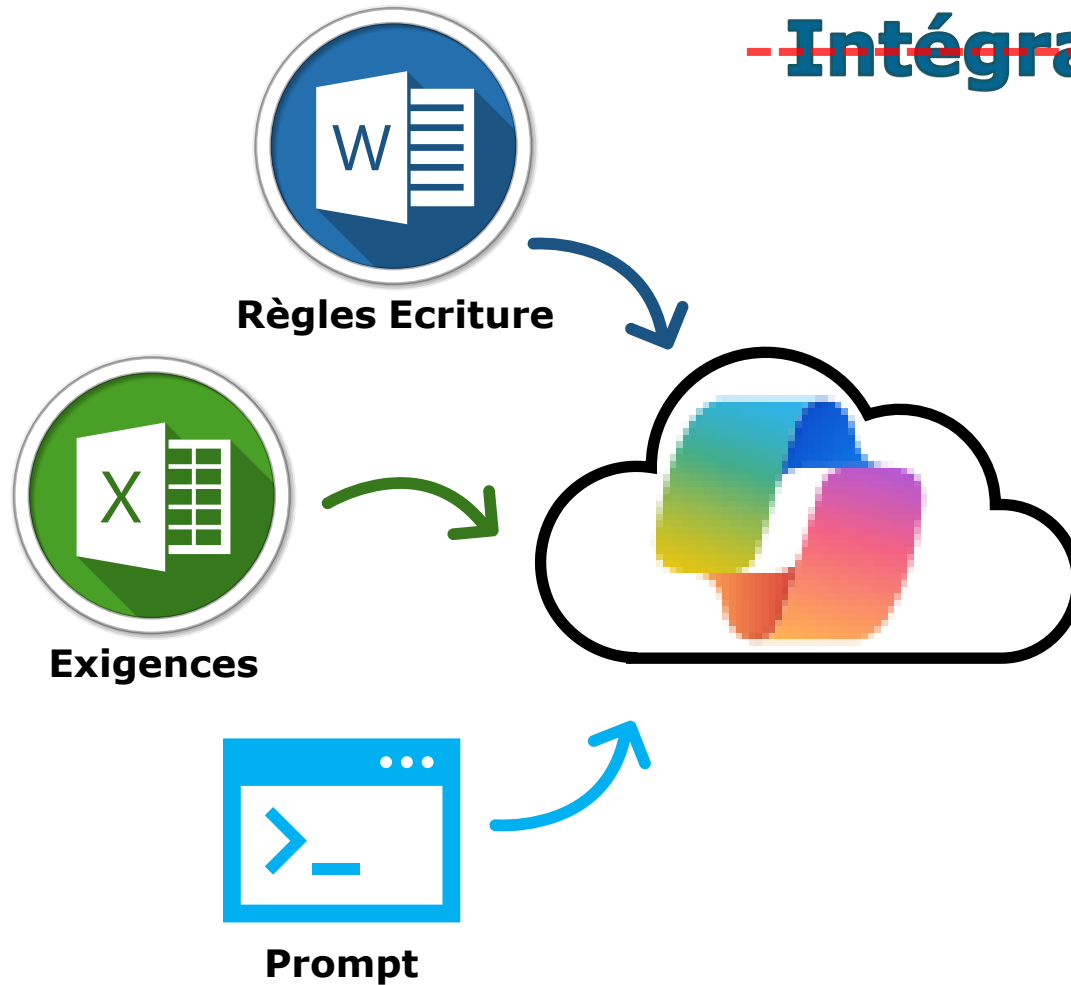
Le Jurrasic Pack.



3. Le Quoi ou les résultats obtenus

Du Big-Bang au Big-Bazar ou Big flop et au lit !

~~Intégration en mode Big-Bang~~

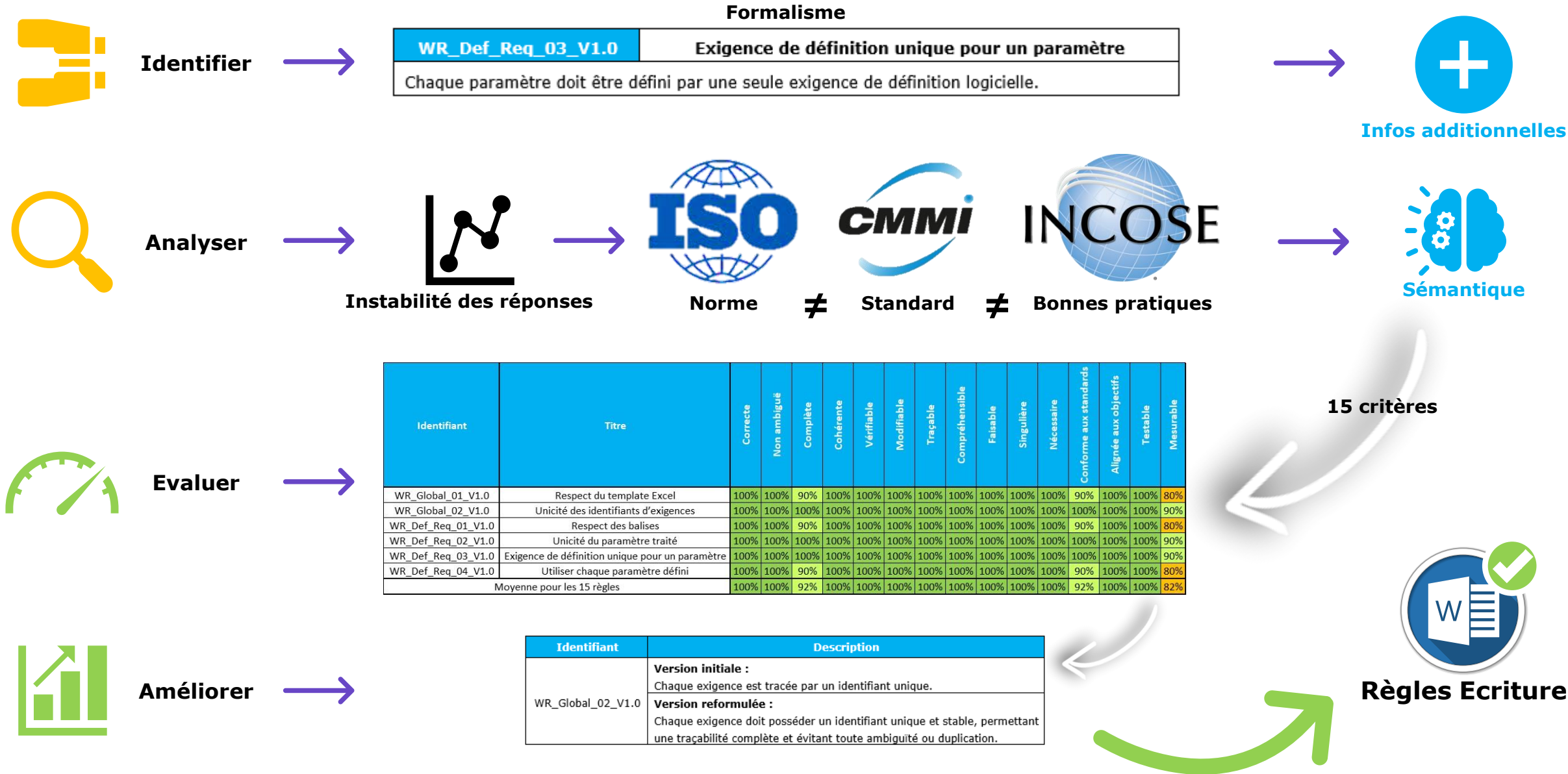


Itératif
Et
Incrémental



3. Le Quoi ou les résultats obtenus

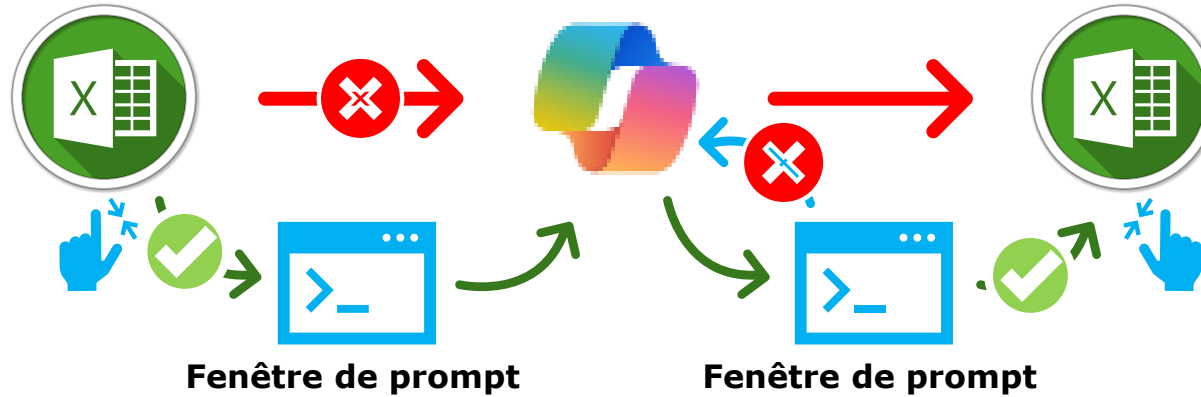
Lundi : gestion des règles d'écriture ?



3. Le Quoi ou les résultats obtenus

Gestion des exigences : tout sauf un long fleuve tranquille.

Des difficultés à l'import de documents et à l'export de résultats



🔍 Pourquoi les résultats dans la fenêtre de prompt sont souvent meilleurs ?

La **fenêtre de prompt** (c'est-à-dire notre échange direct ici) me permet de :

1. **Appliquer une logique humaine et contextuelle** : je peux interpréter les exigences comme le ferait un ingénieur métier, en tenant compte du sens, des intentions, et des conventions fonctionnelles.
2. **Corriger les ambiguïtés** : je peux reformuler intelligemment, ignorer les identifiants techniques (comme les DID), et choisir les bons noms de paramètres.
3. **Adapter dynamiquement** : je peux te poser des questions, ajuster ma méthode, et améliorer la qualité en temps réel.

⚙️ En revanche, dans le traitement automatique (via `python_execution`) :

- Le script est **statique** : il suit des règles prédéfinies, souvent syntaxiques.
- Il ne peut pas **interpréter le sens métier** ou faire des choix qualitatifs.
- Il est limité à ce qu'on lui a codé : pas de jugement, pas d'adaptation.

3. Le Quoi ou les résultats obtenus

Gestion des exigences : Répète un peu pour voir.



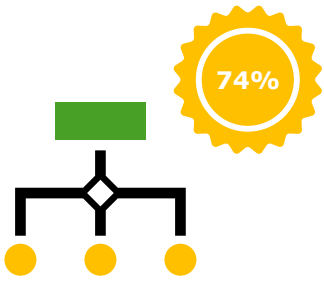
Répétabilité



Exactitude

	Règle 1		Règle 2		Règle 3		Règle 4		Catégorie	
	Répétabilité	Exactitude	Répétabilité	Exactitude	Répétabilité	Exactitude	Répétabilité	Exactitude	Répétabilité	Exactitude
Niveau de confiance Copilot	98%		90%		95%		92%		95%	
Moyenne réelle	100%	100%	99,1%	99,1%	89,7%	86,9%	93,3%	91,0%	93,3%	90,8%
Taux de 100%	100%	100%	96%	96%	48%	46%	78%	77%	77%	76%

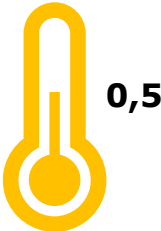
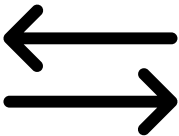
Exigences	Règle 4													
	Essai_1.xlsx	Essai_2.xlsx	Essai_3.xlsx	Essai_4.xlsx	Essai_5.xlsx	Essai_6.xlsx	Essai_7.xlsx	Essai_8bis.xlsx	Essai_9.xlsx	Essai_10.xlsx	Répétabilité	Réalité	Exactitude	
REQ_SYS_S26-FACE_Design_3994(4785)	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX	VRAI	VRAI	VRAI	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_3995(4786)	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX	VRAI	VRAI	VRAI	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_3996(4787)	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX	VRAI	VRAI	VRAI	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_1444(1444)	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	100%	VRAI	100%	
REQ_SYS_S26-FACE_Design_5078(7990)	VRAI	VRAI	FAUX	VRAI	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_1442(1442)	VRAI	VRAI	FAUX	VRAI	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_4026(4826)	VRAI	VRAI	FAUX	VRAI	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_4027(4827)	VRAI	VRAI	FAUX	VRAI	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_5092(8072)	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX	VRAI	VRAI	VRAI	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_715(715)	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	100%	VRAI	100%	
REQ_SYS_S26-FACE_Design_859(859)	FAUX	FAUX	FAUX	FAUX	FAUX	FAUX	FAUX	FAUX	FAUX	FAUX	100%	VRAI	0%	
REQ_SYS_S26-FACE_Design_5670(8376)	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX	VRAI	VRAI	VRAI	70%	VRAI	70%	
REQ_SYS_S26-FACE_Design_771(771)	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	100%	VRAI	100%	
REQ_SYS_S26-FACE_Design_3564(4602)	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	100%	VRAI	100%	



Classification

Regroupement des défauts

Température de l'IA



Règle n°3 : L'exigence contient de la logique combinatoire.



Un prompt rétablissement

3. Le Quoi ou les résultats obtenus

Gestion des 29 exigences de définition logicielle.



MISSION: POSSIBLE



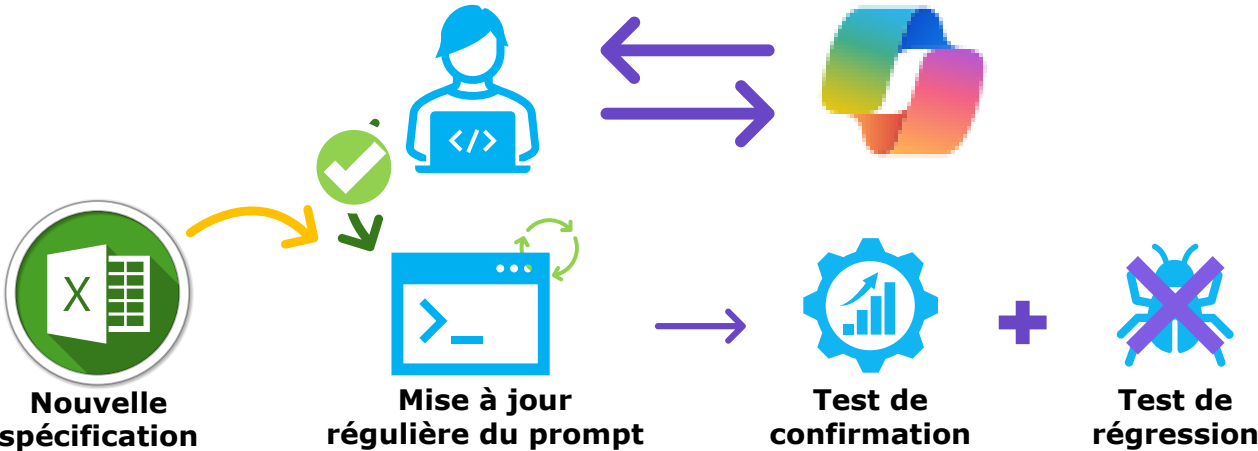
Génération de paramètre fictif

Spécification ADAS

Spécification 2

Spécification 3

Exigences initiales	Exigences formatées Copilot V1	Exigences formatées Copilot V2 (après ajustement)
Target Selection function shall provide the dynamic property of the object (at the target position) 0: Not Defined 1: Unknown 2: Standing (never moved) 3: Not used 4: Crossing left 5: Crossing right 6: Stopped (movable) 7: Moving (movable) 8: Stopped (Oncoming) 9: Moving (Oncoming)	The function Target Selection shall provide the dynamic property of the object through dynamic_property with the following values : 0: Not Defined 1: Unknown 2: Standing (never moved) 3: Not used 4: Crossing left 5: Crossing right 6: Stopped (movable) 7: Moving (movable) 8: Stopped (Oncoming) 9: Moving (Oncoming)	The function Target Selection shall provide the dynamic property of the object (at the target position) through [Parameter name] with the following values : 0: Not Defined 1: Unknown 2: Standing (never moved) 3: Not used 4: Crossing left 5: Crossing right 6: Stopped (movable) 7: Moving (movable) 8: Stopped (Oncoming) 9: Moving (Oncoming)
La fonction RVV consomme le statut de la pédale de frein BRAKE_PEDAL_STATUS défini par : 0 = Released 1 = Pressed avec 0 comme valeur d'initialisation	The function RVV shall receive the brake pedal status through BRAKE_PEDAL_STATUS with the following values : 0 = Released 1 = Pressed with 0 as init value	The function RVV shall receive the brake pedal status through BRAKE_PEDAL_STATUS with the following values : - 0 = Released - 1 = Pressed with 0 as init value
[Ctrl_btDsThr] : flux de contrôle pour l'information Tair_admission [Ctrl_btDsThr] est de type : FLUXFONCTIONNEL [Ctrl_btDsThr] prend la valeur : 0 : Inactif à l'initialisation Fonction consommatrice : Acq_tDsThr ; [Ctrl_btDsThr] possède la résolution suivante : 1 [Ctrl_btDsThr] est produite avec la précision suivante : - [Ctrl_btDsThr] est bornée à la plage suivante : [0 1] 0 : Inactif 1 : Actif [Ctrl_btDsThr] est précalibrée à la valeur -	The function Acq_tDsThr shall receive control flow for intake air temperature information through Ctrl_btDsThr with the following values : 0 : Inactif 1 : Actif with 0 as init value	The function Acq_tDsThr shall receive control flow for Tair_admission through Ctrl_btDsThr with the following values : 0 : Inactif 1 : Actif with 0 as init value



3. Le Quoi ou les résultats obtenus

Gestion des 57 exigences fonctionnelles.



Simplification combinatoire de 6 exigences

Exigences initiales	Exigences formatées Copilot	Proposition de simplification combinatoire de Copilot	Explications de Copilot
IF OMS_OEM_MODE_STATUS = Disabled Then DRIVER_FATIGUE_ALERT_STATUS_UNFILTERED = No Alert DRIVER_DISTRACTION_ALERT_STATUS_UNFILTERED = No Alert OMS_CANCEL_REASON_INTENRAL =No display INTERIOR_CAMERA_CANCEL_REASON_INTERNAL=No display OMS_STATUS_DISPLAY_UNFILTERED=No display OMS_DRIVER_GUIDE_UNFILTERED=No display ELSE if OMS_OEM_MODE_STATUS = Not disabled then No impact on alert management	IF0 (OMS_OEM_MODE_STATUS = Disabled) THEN (DRIVER_FATIGUE_ALERT_STATUS_UNFILTERED = No Alert DRIVER_DISTRACTION_ALERT_STATUS_UNFILTERED = No Alert OMS_CANCEL_REASON_INTENRAL = No display INTERIOR_CAMERA_CANCEL_REASON_INTERNAL = No display OMS_STATUS_DISPLAY_UNFILTERED = No display OMS_DRIVER_GUIDE_UNFILTERED = No display) ELSEIF0 (OMS_OEM_MODE_STATUS = Not disable d) THEN (No impact on alert management)	IF (OMS_OEM_MODE_STATUS = Disabled) THEN (DRIVER_FATIGUE_ALERT_STATUS_UNFILTERED = No Alert AND DRIVER_DISTRACTION_ALERT_STATUS_UNFILTERED = No Alert AND OMS_CANCEL_REASON_INTERNAL = No display AND INTERIOR_CAMERA_CANCEL_REASON_INTERNAL = No display AND OMS_STATUS_DISPLAY_UNFILTERED = No display AND OMS_DRIVER_GUIDE_UNFILTERED = No display) ELSE (No impact on alert management)	Modifications apportées : Correction de la faute de frappe dans OMS_CANCEL_REASON_INTENRAL → OMS_CANCEL_REASON_INTERNAL Fusion des blocs IF0 / ELSEIF0 en IF / ELSE pour plus de clarté Remplacement de ELSEIF0 (OMS_OEM_MODE_STATUS = Not disabled) par un simple ELSE, car cela couvre logiquement tous les autres cas Respect du formalisme IF / THEN / ELSE Aucune modification des noms de variables ou des valeurs (hors correction de typo)



- **La présence de multiples "and" au sein d'une même phrase.**
"If driver face is present *and* eyes opened in the FOV *and* IQ Volume DMC"
- **La présence d'informations additionnelles entre parenthèses.**
"if driver's eyes are not_present_in_the_the_FOV_and_IQ_Volume_DMC
(*DRIVER_EYES_ON_ROAD_STATUS = Unavailable*)"

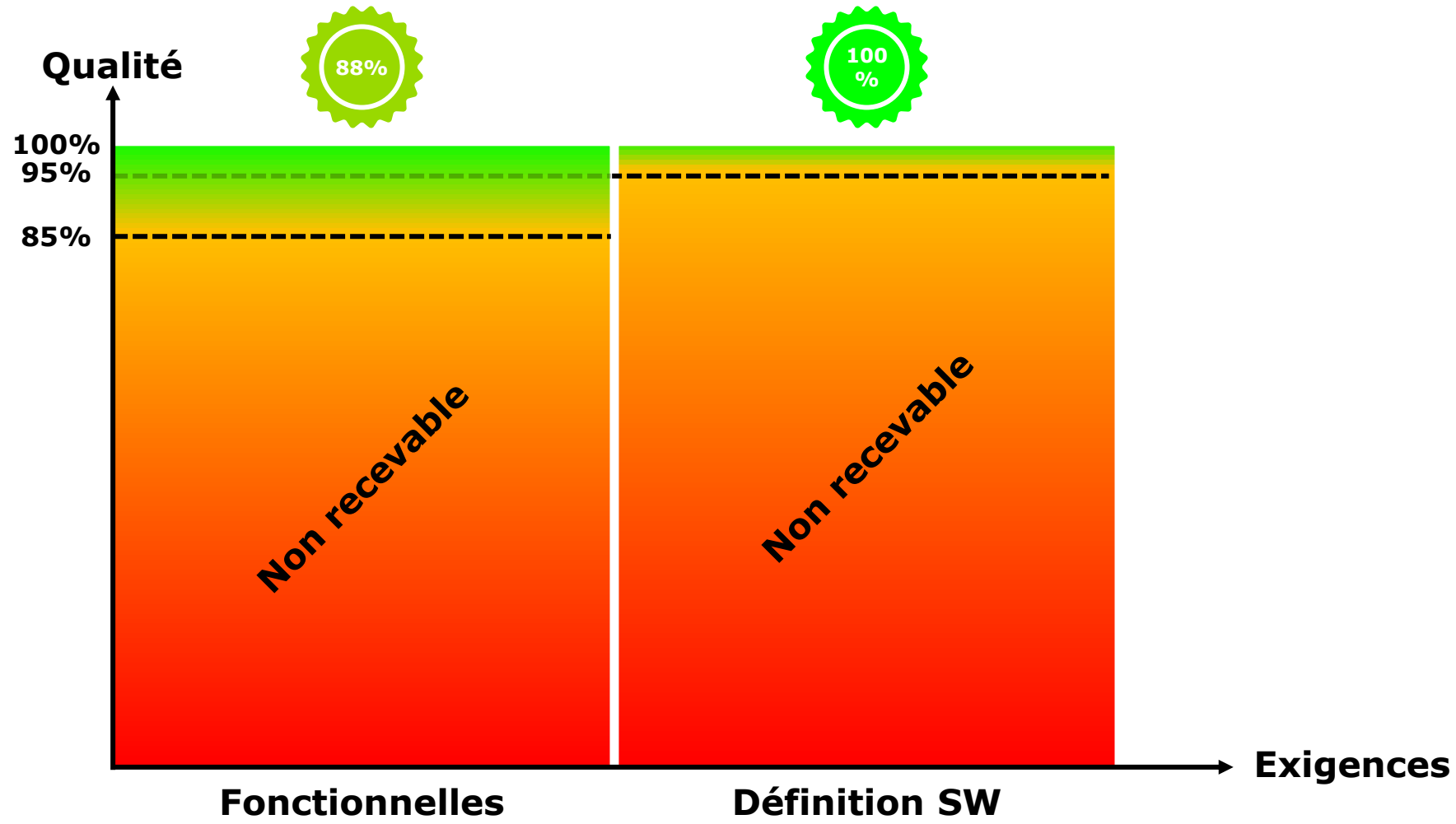
- Des premiers résultats très probants :**
- Identification d'une condition ELSEIF1 inutile.
 - Simplification d'une recopie de condition dans un paramètre.
 - Identification d'un doublon d'affectation (erreur de Copier/Coller).
 - Regroupement de 2 blocs IF successifs en 1 seul avec la logique AND.
 - Prise en compte des priorités entre AND et OR.



- Approximations :**
- Supposition que tous les paramètres sont booléens.
 - Si manque d'informations complémentaires.

4. L'heure du bilan

La qualité avant tout.



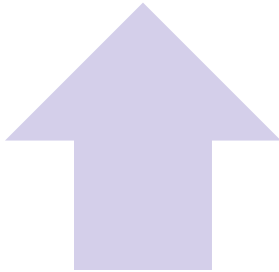
4. L'heure du bilan

Le temps est infini contrairement aux délais.

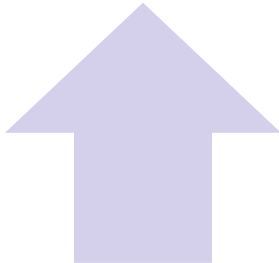


productivité

+25%



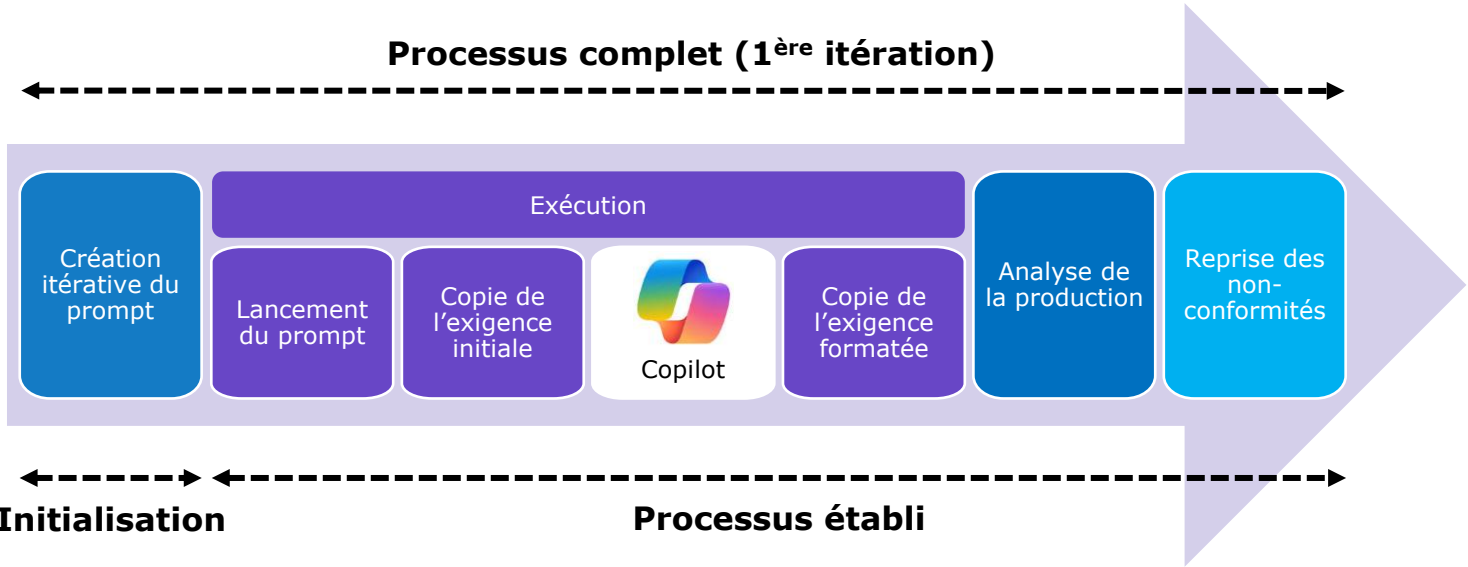
+92%



30h

en mode manuel

vs



Estimation temporelle								
Fixe		20h	10s	-	-	-	-	-
Fonction du nombre d'exigence	Définition SW	-	-	10s	10s	10s	20s	1min
	Fonctionnelle	-	-				60s	5min
Total cas client 1 29 exigences définition SW 57 exigences fonctionnelles		20h	10s	14min20s	14min20s	14min20s	1h06min40s	35min
			2h24min50s					


4. L'heure du bilan

Une solution vraiment dans le coût ?



Profil junior

+124%
pour la 1^{ère} itération



Profil sénior

Estimation
d'amortissement
pour le cas client

116

Exigences supplémentaires
traitées par un profil junior

50% d'évolution
entre Vn et Vn+1

Version de spécification	V1	V2	V3	V4	V5
Nombre estimé d'évolution d'exigences	86	43	21	10	5
Total d'exigences à (re)traiter	86	79			



+40%

4. L'heure du bilan

Chaque problème a une solution ...

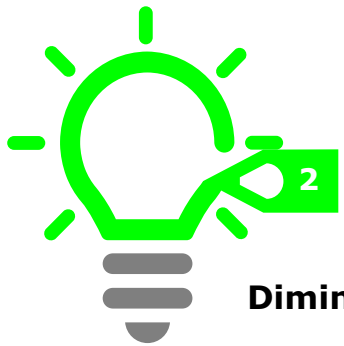


Amortissement sur d'autres spécifications.

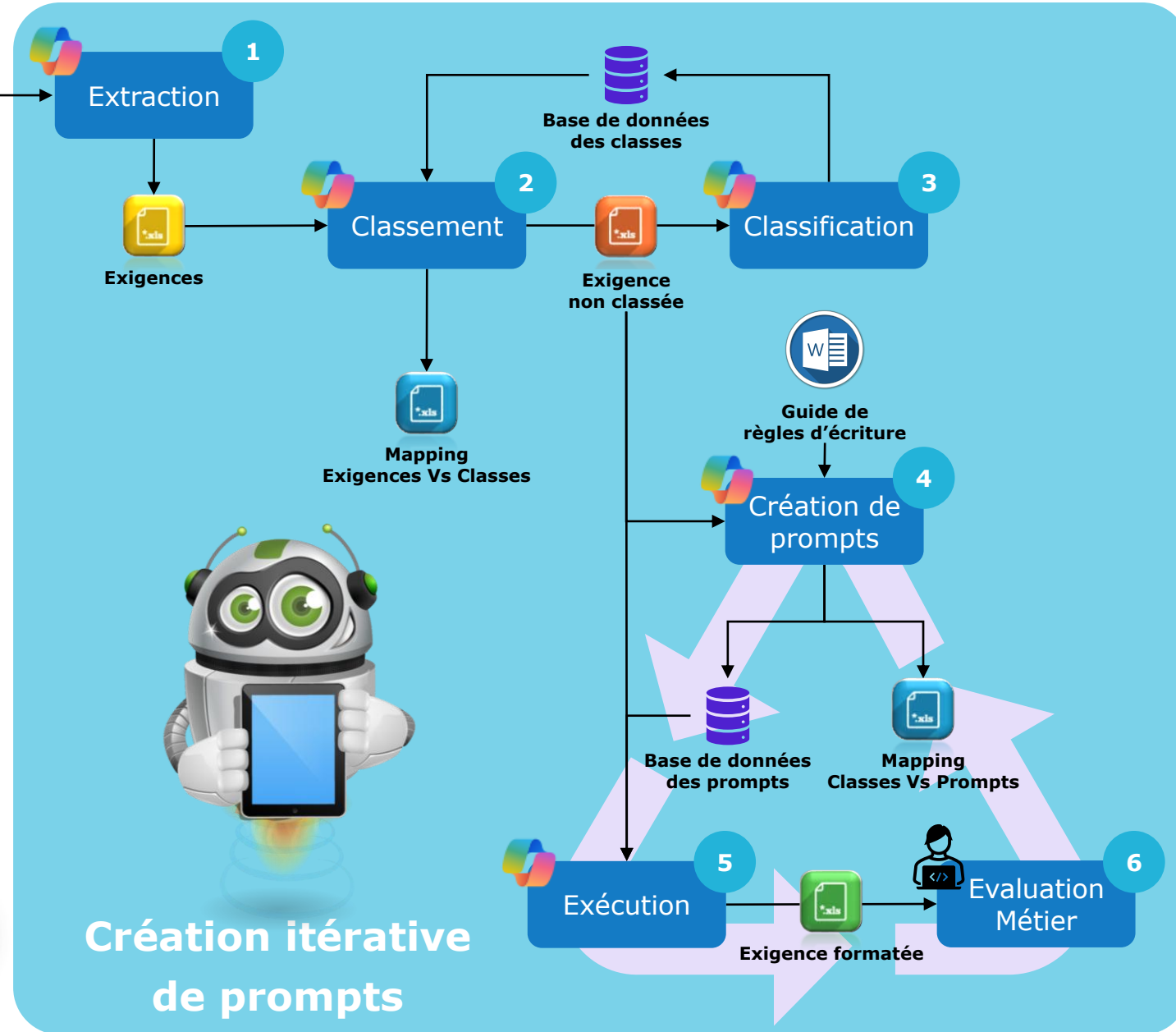


Quid de la genericité du prompt pour les exigences fonctionnelles ?

OU / ET

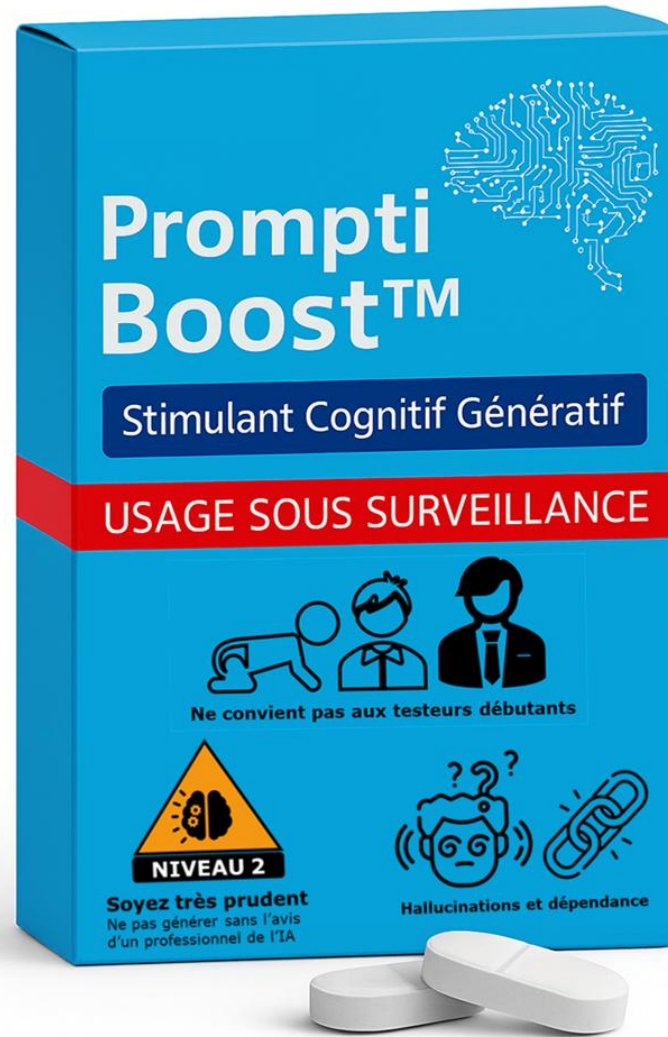


Diminuer le coût de la 1^{ère} itération.



même si celle-ci n'est pas forcément celle que l'on espérait ou attendait initialement (Alice Hoffman).

4. L'heure du bilan



LES IA GENERATIVES

C'EST PUISSANT... MAIS PAS OMNISCIENT !



Utilisez-les quand il faut et uniquement quand il faut !

LES IA GENERATIVES NE REMPLACENT PAS L'EXPERTISE HUMAINE, LA CONNAISSANCE METIER OU LES PROCESSUS DE TEST

4. L'heure des interrogations

Les défis de l'IA au quotidien



Maintenance des prompts



Gestion des compétences à moyen terme

CEDRIC

LEFEBVRE

Applicabilité de l'IA générative
pour le test des systèmes
embarqués

(focus sur l'aide à l'analyse des exigences)

Merci pour votre attention.



JOURNÉE
THÉMATIQUE
IA GÉNÉRATIVE
POUR TESTER

[expleo]
Think bold, act reliable