

Comparer la même demande entre deux assistants

Mesurer les ajouts sur un cas fictif sans refaire le module C2

Module CL2 - FAIRE - 5 min - version 1.1

C2 t'a déjà appris à rédiger. CL2 n'a donc de valeur que s'il t'apprend autre chose : comparer deux sorties contre la même matière et documenter ce que chacune a ajouté.

Un seul cas ne permet pas de dire qu'un outil « a l'habitude » de faire X. Il permet seulement de dire : sur cette exécution, voici les ajouts observés.

Ce que tu sauras faire en sortant

- Rejouer exactement la même demande et la même matière fictive.
- Produire un tableau sortie élément ajouté type action.
- Vérifier les ajouts toi-même au lieu de faire confiance à l'auto-critique du modèle.
- Refuser de transformer trois essais en classement général des assistants.

L'essentiel, en trente secondes

Prépare une matière fictive avec quelques absences volontaires :

`text Client fictif : Atelier Martin Objet : relance d'une facture fictive Montant : NON FOURNI Date de règlement : NON FOURNIE Fait fourni : facture arrivée à échéance Ton : posé `

Demande identique dans les deux outils :

`text Rédige une relance courte uniquement à partir de la matière fournie. N'ajoute aucun montant, date, pénalité, article, délai ou engagement. Si une information manque, ne la complète pas. `

Puis compare chaque sortie à la matière, pas d'abord à l'autre sortie.

Tableau de contrôle

Outil / exécution	Élément de sortie	Présent dans la matière ?	Type	Action	--- --- ---: --- ---	A	« règlement sous 8 jours »	non	délai ajouté	retirer	B	formule de politesse	non	style	décider selon préférence	B	aucun chiffre ajouté	-	constat du cas	aucune généralisation
-------------------	-------------------	---------------------------	------	--------	----------------------	---	----------------------------	-----	--------------	---------	---	----------------------	-----	-------	--------------------------	---	----------------------	---	----------------	-----------------------

Types possibles : fait, chiffre, date, engagement, référence, style, structure.

Un ajout de style n'a pas le même risque qu'un montant ou une référence. Le tableau rend cette différence visible.

Micro-défi

La sortie Claude contient une formule de politesse non demandée. L'autre outil ajoute un délai de 8 jours. Que peux-tu conclure ?

Attendu : sur ce cas, Claude a ajouté un élément de style et l'autre outil un engagement temporel. Tu ne peux pas conclure que Claude « ajoute toujours du style » ni que l'autre « invente toujours des délais ».

Auto-contrôle de l'outil : utile mais insuffisant

Tu peux demander :

`text Liste les éléments de ta réponse qui n'étaient pas explicitement fournis dans ma matière. Ne les corrige pas. `

Mais cette liste est une piste. Le même système relit sa propre production et peut manquer un ajout. Le contrôle de référence reste ta comparaison matière sortie.

Le cas où ça se passe mal

Cas reconstitué.

Trois exécutions fictives montrent le même type d'ajout. Tu écris dans ta fiche : « Claude fait toujours X ».

Le problème est la portée : trois observations ne démontrent pas un comportement universel, et une mise à jour du produit peut changer le résultat.

Correction : écris observé dans N exécutions datées sur ce cas, et transforme l'observation en point de surveillance, pas en doctrine.

Corriger le prompt ou simplement retirer ?

Si un ajout important se reproduit dans tes propres tests, tu peux ajouter une consigne ciblée au prompt et retester.

Mais ne déduis pas :

- qu'un prompt corrigé est « validé » sans replay documenté ;
- qu'une consigne empêche définitivement l'ajout ;
- qu'un comportement observé sur relance s'applique à un devis ou à un autre type de texte.

La correction est locale au cas testé.

Ce que ce module ne règle pas

- Rédaction métier de C2.
- Contenu contractuel d'un devis.
- Comparatif général Claude/ChatGPT.
- Classement de qualité, vitesse ou fiabilité des assistants.
- « Habitudes » stables d'un outil sans banc de test dédié.
- Données réelles : l'exercice reste fictif.

Quand tu t'arrêtes

- Tu modifies la demande entre les deux outils : comparaison invalide.
- Tu utilises un document ou une facture réelle pour tester.
- Tu écris toujours, meilleur, plus fiable à partir de quelques exécutions.
- Une mention importante n'a pas de provenance dans la matière.

À toi

- Crée un cas fictif avec deux informations volontairement absentes.

- Rejoue exactement la même demande dans deux assistants.
- Complète le tableau de contrôle.
- Demande l'auto-liste des ajouts et compare-la à ton propre relevé.
- Écris deux lignes : observé / non démontré.

C'est réussi si tu peux documenter les différences du cas sans les transformer en caractéristiques générales des outils.

Vérifie ton geste

- Le texte est plus fluide. Ce n'est ni un défaut ni une preuve de qualité ; contrôle les faits ajoutés.
- Auto-contrôle dit « aucun ajout ». Compare quand même à la matière.
- Même ajout sur trois essais. Point de surveillance local, pas toujours.
- Un article juridique apparaît sans matière. Retirer ; ne pas le traiter comme piste de droit dans cet exercice.
- Facture réelle disponible. Ne pas l'utiliser pour le banc ; cas fictif suffisant.

Seuil de réussite : 5 sur 5 sur la méthode comparative.

Selon ce que tu obtiens

- Ajouts différents documente-les selon leur risque.
- Aucun ajout factuel bon résultat pour ce cas, pas certificat de l'outil.
- Auto-contrôle manque un élément preuve utile des limites de l'auto-relecture.
- Bloqué compare une seule phrase générée à quatre faits source.

Fiche mémo

- Même matière, même prompt.
- Comparer chaque sortie à la matière.
- Ajout = type + action.
- Auto-contrôle = piste, pas preuve.
- N essais datés comportement universel.
- Prompt modifié = NON REJOUÉ jusqu'au nouveau test.

Ce que tu viens de pratiquer

- CL2 n'est plus un miroir de C2 : il enseigne une comparaison de sorties.
- Tu as remplacé les « habitudes de l'outil » par des observations datées.
- Tu sais séparer ajout de style et ajout factuel engageant.

Sources et statut : docs/bloc-claude-PROVENANCE.md`. CL2 reste S2 et brouillon et dépend des relevés produit de CL1.