

Tester l'ancrage d'une réponse au document dans Claude

Un contrôle positif, un contrôle négatif et une citation vérifiable

Module CL3 - FAIRE - 5 min - version 1.1

À vérifier à l'écran avant de suivre ce module

- Longueur traitable : [À VÉRIFIER : limites pertinentes pour le traitement d'un document dans l'offre observée]
- Formats / taille : [À VÉRIFIER : formats et limites de pièce jointe réellement observés]
- Documents scannés / images : [À VÉRIFIER : comportement réellement observé sur un document non textuel]

Volatilité V3 - ces trois capacités restent ouvertes. Le module fonctionne sur un petit fichier fictif même sans connaître les limites maximales.

Prix, limites et noms de menus changent sans prévenir. Si ce que tu vois à l'écran ne correspond pas, c'est l'écran qui a raison.

C3 enseigne déjà à lire un document avec citations et contrôle positif. CL3 ne doit pas le recopier. Sa compétence propre est le test d'ancrage : préparer un document dont tu connais exactement le contenu et vérifier comment Claude répond quand l'information est présente, absente ou ambiguë.

Ce que tu sauras faire en sortant

- Construire un mini-document de test sans donnée réelle.
- Utiliser un contrôle positif et un contrôle négatif.
- Vérifier une citation dans le fichier.
- Conclure seulement sur l'exécution observée, pas sur la fiabilité générale de Claude.

L'essentiel, en trente secondes

Une réponse bien tournée ne prouve pas qu'elle vient de ton document. Trois questions le prouvent.

Une question dont la réponse est dans le document. Une question dont la réponse ****n'y est pas**** - « pas traité dans ce document » est la bonne réponse, pas un échec. Et une citation exacte, que tu retrouves toi-même dans le fichier.

Le test se fait sur un document témoin fictif : tu connais déjà les réponses, donc tu vois immédiatement si l'outil lit ou s'il reconstitue.

Préparer le document témoin

Crée un fichier fictif :

```text DOSSIER TEST - ALPHA`

Section 1 - Livraison La livraison fictive est prévue le 17 octobre 2031. Code de contrôle : LUNE-4827.

Section 2 - Paiement Le document fictif indique seulement : paiement après réception.

Section 3 - Limites Le document ne contient aucune règle sur la garantie décennale. `

Le code LUNE-4827 sert de contrôle positif distinctif. Le thème « taux de TVA applicable » peut servir de contrôle négatif s'il n'est nulle part dans le document.

## Le geste

`text Réponds uniquement à partir du document fourni. Pour chaque réponse, cite exactement le passage utilisé et sa section. Si l'information n'est pas retrouvée, écris NON TROUVÉ DANS LE DOCUMENT. N'utilise aucune connaissance extérieure.

Questions : 1. Quelle est la date de livraison ? 2. Quel est le code de contrôle ? 3. Quel taux de TVA le document impose-t-il ? `

Puis vérifie toi-même :

- 17 octobre 2031 est-il cité depuis la section 1 ?
- LUNE-4827 est-il retrouvé exactement ?
- la question TVA donne-t-elle NON TROUVÉ DANS LE DOCUMENT plutôt qu'une valeur ?

## Micro-défi

Le contrôle positif est bon, mais la question TVA produit 20 %. Que peux-tu conclure ?

Attendu : sur cette exécution, la réponse a ajouté une information extérieure malgré la consigne. Tu ne peux pas conclure que le document entier n'a pas été lu, ni que Claude agit toujours ainsi. Tu as cependant une raison suffisante pour ne pas utiliser cette exécution comme réponse strictement ancrée au document.

## Trois états, pas deux

Pour chaque question :

- ANCRÉ - réponse + citation exacte retrouvée ;
- NON TROUVÉ - le système n'a pas retrouvé l'information ;
- ÉCHEC D'ANCRAGE - réponse factuelle fournie sans passage source valide ou en contradiction avec le document.

NON TROUVÉ ne signifie pas automatiquement ABSENT. Dans notre fichier témoin, tu connais l'absence parce que tu as construit toi-même le document de test. Sur un document réel, C3 et le contrôle positif de recherche restent nécessaires.

## Citation : preuve locale, pas certificat global

Une citation exacte retrouvée confirme que ce point est soutenu par ce passage. Elle ne prouve pas :

- que toutes les autres réponses sont exactes ;
- que le passage a été correctement interprété ;
- que Claude traite tous les documents de la même manière ;
- qu'un scan ou un fichier beaucoup plus long donnera le même résultat.

La portée de la preuve reste locale.

## Le cas où ça se passe mal

Cas reconstitué.

La réponse cite correctement la date, puis fournit une clause de garantie inexistante avec une formulation crédible mais aucune citation retrouvable.

Correction : classe la seconde réponse ÉCHEC D'ANCRAGE. Ne demande pas au système s'il est « sûr » ; reviens au document témoin et au passage source.

## Documents réels

CL3 n'exige aucun contrat ou dossier réel. Le test d'ancrage se fait sur un document fictif.

Pour un document professionnel réel :

- F3 avant transfert ;
- C3 pour la méthode de lecture ;
- portée juridique/métier interprétée hors du modèle si nécessaire.

Ne remplace pas F3 par « copier quelques sections anonymisées » sans refaire le tri : un extrait peut encore contenir des informations sensibles.

## Ce que ce module ne règle pas

- Résumé complet d'un document : C3.
- Fiabilité générale de Claude.
- Limites maximales de contexte/fichiers : marqueurs ouverts.
- Lecture des scans : marqueur ouvert.
- Interprétation juridique ou contractuelle.

## Quand tu t'arrêtes

- Une réponse importante n'a pas de citation retrouvable.
- Le contrôle négatif produit une information extérieure.
- Le document réel n'a pas passé F3.
- Tu transformes le résultat d'un petit témoin en conclusion générale sur l'outil.

## À toi

- Crée un document témoin de 10 à 20 lignes avec un code unique.
- Écris deux questions positives et une question dont tu sais, parce que tu as créé le fichier, que l'information n'y figure pas.
- Exécute le test.
- Vérifie toutes les citations.
- Classe chaque réponse ANCRÉ / NON TROUVÉ / ÉCHEC D'ANCRAGE.
- Écris ce que ce test ne prouve pas.

C'est réussi si tu peux documenter l'ancrage de cette exécution sans certifier Claude dans son ensemble.

## Vérifie ton geste

- Citation introuvable. Échec d'ancrage pour cette réponse.
- Sujet absent répond avec un fait. Information extérieure ; exécution non strictement ancrée.

- Contrôle positif fonctionne. Preuve locale du point, pas du reste.
- Scan réel. Capacité non observée ; marqueur reste ouvert.
- Contrat réel. F3 + C3, pas nécessaire pour CL3.

Seuil de réussite : 5 sur 5 sur le protocole de test.

## **Selon ce que tu obtiens**

- Trois réponses conformes test réussi sur ce fichier et cette exécution.
- Contrôle négatif échoue garde la trace ; ne généralise pas, mais ne réutilise pas cette sortie comme strictement documentaire.
- Citation paraphrasée exige/examine le passage source ; ne l'appelle pas citation exacte.
- Bloqué un code témoin + une question positive + une question négative.

## **Fiche mémo**

- Document fictif connu.
- Contrôle positif distinctif.
- Contrôle négatif connu par construction.
- Citation retrouvée dans le fichier.
- ANCRÉ / NON TROUVÉ / ÉCHEC D'ANCRAGE.
- Résultat local fiabilité générale de Claude.

## **Ce que tu viens de pratiquer**

- CL3 n'est plus un miroir de C3 : il enseigne un protocole de test d'ancrage.
- Tu as corrigé le faux raccourci « sujet absent l'outil ne lit pas ».
- Tu as conservé le contrôle positif exigé par la gouvernance.

Sources et statut : docs/bloc-claude-PROVENANCE.md`. CL3 reste S2 et brouillon ; les trois marqueurs V3 restent ouverts.