NL-Q · FICHE-MÉMO

Avant qu'une IA écrive dans vos données : l'arbre du retour arrière

Ce que l'écriture apporte.
L'assistant ne se contente plus de répondre : il met à jour un tableur, corrige des fiches dans votre base, saisit dans votre logiciel de gestion.

Ce qu'elle ouvre.
Une modification fausse ne se voit pas toujours le jour même. Sans copie ni trace, on ne sait ni la défaire, ni dire ce qui a changé.

Ce que contient la fiche.
Cinq questions à poser dans l'ordre, une issue pour chaque réponse, puis huit questions pour vérifier.

Pour les dirigeants et responsables de TPE et PME, non développeurs, qui s'apprêtent à laisser un assistant IA modifier leurs données

Édition du 19/11/2026 · NLQ

01 Suivez les cinq questions dans l'ordre

Une réponse fausse se corrige en la relisant ; une écriture fausse reste dans vos données, et sert ensuite à un devis, une relance, une commande. L'OWASP le souligne : qu'un assistant se trompe seul ou qu'il suive une consigne piégée, les dégâts dépendent de ce qu'il a le droit de faire, quelle qu'en soit la cause.

1 **Écrire ?** Nécessaire ou pas → 2 **Sauvegarde ?** Déjà restaurée → 3 **Copie ?** Essai à part → 4 **Trace ?** Avant et après → 5 **Validation ?** Un nom

Chaque réponse mène à une issue. Un « oui » à la cinquième ne rattrape pas un « non » à la deuxième.

QUESTION	SI NON	SI OUI
1. L'écriture est-elle nécessaire ?	L'assistant propose, une personne saisit	Question 2
2. Une sauvegarde récente a-t-elle déjà été restaurée ?	Arrêt : sauvegarder et restaurer une fois d'abord	Question 3
3. Peut-on travailler sur une copie ?	Premier lot réduit à quelques fiches notées avant, puis question 4	Essai sur la copie, puis question 4
4. Chaque modification est-elle tracée ?	Une personne relève l'état avant et après chaque lot, sinon arrêt	Question 5
5. Une personne nommée valide-t-elle ?	Pas d'écriture	Écriture sur l'original, lot par lot

02 Demandez d'abord si l'écriture est nécessaire

Beaucoup de demandes d'écriture sont en fait des propositions : une liste de doublons à fusionner, des prix à revoir, des fiches à compléter. L'OWASP cite le cas d'un outil choisi pour lire des documents qui sait aussi les modifier et les supprimer : une fonction dont personne n'avait besoin, mais que l'assistant peut utiliser.

La question à poser.
« Que doit-il changer, exactement, et pourquoi ne peut-il pas seulement me le proposer ? » Si la réponse tient en une liste, une proposition suffit.

La proposition dans un fichier à part.
L'assistant rend un tableau « ligne, valeur actuelle, valeur proposée ». Une personne le relit et saisit : vos données ne changent que par une main connue.

Une action nommée, pas un fichier entier.
S'il doit écrire, on nomme l'action : « mettre à jour la colonne Statut », pas « gérer le fichier clients ». L'OWASP recommande de limiter au strict nécessaire les fonctions offertes à l'assistant.

03 Exigez une sauvegarde déjà restaurée

Une sauvegarde qu'on n'a jamais restaurée n'a pas fait ses preuves. Cybermalveillance.gouv.fr le rappelle : il arrive que la sauvegarde ne se fasse pas correctement, d'où le conseil de vérifier régulièrement qu'elle fonctionne, par exemple en la recopiant dans le système d'origine.

Une sauvegarde juste avant.
Pour un fichier souvent mis à jour, c'est sa version la plus récente qu'il faut pouvoir retrouver (Cybermalveillance.gouv.fr). On sauvegarde donc le jour de l'écriture, en plus du calendrier habituel.

Un essai de restauration.
Restaurez une fois, dans un dossier d'essai, et ouvrez le résultat : le tableur s'ouvre-t-il, les dernières lignes saisies y sont-elles ?

Une copie hors d'atteinte.
Débranchez le support après la sauvegarde et gardez-en une copie dans un autre lieu, comme le conseille Cybermalveillance.gouv.fr. Vérifiez aussi que le compte de l'assistant n'a pas accès à la sauvegarde.

04 Essayez sur une copie avant l'original

Une copie du tableur, ou un espace d'essai de votre logiciel s'il en propose un, permet de voir l'effet de l'écriture sans rien risquer. On compare ensuite la copie modifiée à l'original : ce qui diffère, c'est exactement ce que l'assistant a fait.

Copie d'essai	Sauvegarde restaurée	Historique de l'outil	Relevé à la main	Aucun moyen
rien à défaire	retour à un état connu	s'il existe	partiel	on ne sait pas

← MOINS DE RISQUE PLUS DE RISQUE →

La copie, puis la comparaison.
Avant de valider, faites lister les lignes qui diffèrent entre la copie modifiée et l'original, avec l'ancienne et la nouvelle valeur.

Une copie qui ressemble au vrai.
Une copie réduite à quelques lignes propres ne montre pas les cas qui piègent : cellules vides, doublons, noms accentués, dates mal saisies.

Sans copie possible, un petit lot.
Si le logiciel n'offre pas d'essai, la première écriture porte sur quelques fiches choisies, dont on a noté l'état avant.

05 Gardez la trace de ce qui a changé

Le jour où un chiffre paraît faux, trois questions se posent : qui l'a changé, quand, et que valait-il avant ? L'ANSSI recommande de journaliser l'ensemble des traitements de l'assistant, dont ses appels à des outils et à des données, pour pouvoir reconstituer entièrement un événement (R29). Pour l'OWASP, surveiller l'activité de l'assistant et des systèmes qu'il touche n'empêche pas l'erreur, mais en limite l'étendue.

Cinq informations par modification.
Le compte qui écrit, la date et l'heure, l'endroit (fichier, table, ligne), la valeur avant, la valeur après. C'est la liste à demander au prestataire.

La demande d'origine aussi.
L'ANSSI range les requêtes des utilisateurs parmi ce qu'il faut journaliser, en protégeant les données sensibles qu'elles contiennent.

Un rythme limité.
L'OWASP propose de limiter le nombre d'actions possibles sur une période, pour repérer une dérive avant qu'elle ne s'étende. En pratique : des lots, pas toute la base d'un coup.

Sans trace, un relevé.
Si l'outil ne garde rien, une personne exporte l'état avant et après chaque lot. Si personne ne peut le faire, l'écriture attend.

06 Nommez la personne qui valide

L'ANSSI recommande qu'un système d'IA ne puisse pas exécuter seul des actions critiques, comme une transaction bancaire ou ce qui touche directement des personnes, et qu'une décision à fort impact passe par un contrôle humain (R9). L'OWASP recommande l'accord d'une personne avant les actions à fort impact, et que le droit d'agir soit vérifié par le logiciel qui reçoit l'écriture, pas laissé au jugement de l'assistant.

Un nom par type de modification.
La personne qui connaît les données valide : le responsable des stocks pour les quantités, la personne qui tient la facturation pour les fiches clients.

Elle valide sur la liste des changements.
On lui montre les lignes qui changent, avant et après, pas un résumé rédigé par l'assistant.

Rien ne part d'un mail reçu.
Une écriture déclenchée par un mail ou une page web attend une validation : l'ANSSI recommande de limiter, voire de proscrire, les actions automatiques à partir d'entrées non maîtrisées (R27).

Comment savoir si c'est en place

Huit questions fermées. Chaque « non » renvoie à une question de l'arbre.

- [ ] Savez-vous dire, en une phrase, ce que l'assistant a le droit de modifier ?
- [ ] Avez-vous déjà restauré une sauvegarde et ouvert le résultat ?
- [ ] La première écriture a-t-elle eu lieu sur une copie ou sur un petit lot noté avant ?
- [ ] Chaque lot a-t-il été validé par une personne nommée, sur la liste des lignes changées ?
- [ ] La dernière sauvegarde date-t-elle d'avant la dernière écriture de l'assistant ?
- [ ] Le compte de l'assistant est-il privé d'accès à la sauvegarde ?
- [ ] Pouvez-vous retrouver, pour une modification d'hier, la valeur d'avant ?
- [ ] Toute écriture née d'un mail ou d'une page web attend-elle une validation ?

VOTRE PROCHAINE ÉTAPE

Venez avec la première modification que vous voulez confier à votre assistant : on la fait passer ensemble dans les cinq questions.

Réserver un échange gratuit

www.nl-q.com/rendez-vous/

Sources et portée

Toutes les sources ont été ouvertes le 19/11/2026.

- OWASP Gen AI Security Project, *LLM06:2025 Excessive Agency* : fonctions excessives, accord humain, journal, rythme limité. https://genai.owasp.org/llmrisk/llm062025-excessive-agency/ (consulté le 19/11/2026)
- ANSSI, guide ANSSI-PA-102 *Recommandations de sécurité pour un système d'IA générative*, version 1.0 du 29/04/2024 : R9, R27, R29. https://messervices.cyber.gouv.fr/documents-guides/Recommandations_de_s%C3%A9curit%C3%A9_pour_un_syst%C3%A8me_d_IA_g%C3%A9n%C3%A9rative.pdf (consulté le 19/11/2026)
- Cybermalveillance.gouv.fr, *Pourquoi et comment bien gérer ses sauvegardes ?* : version récente, support débranché, autre lieu, test. https://www.cybermalveillance.gouv.fr/tous-nos-contenus/bonnes-pratiques/sauvegardes (consulté le 19/11/2026)
- Cybermalveillance.gouv.fr, *Les 10 mesures essentielles pour votre cybersécurité*, mesure 2. https://www.cybermalveillance.gouv.fr/tous-nos-contenus/bonnes-pratiques/10-mesures-essentielles-assurer-securite-numerique (consulté le 19/11/2026)

Portée : l'ordre des questions est proposé par cette fiche ; les exemples sont fictifs. Ni les obligations liées aux données personnelles ni la tenue de la comptabilité ne sont traitées. Ni audit, ni avis juridique.

Édité par NLQ — NL-Q (www.nl-q.com), avec le soutien d'ALMA, sponsor. Édition du 19/11/2026. Exemples fictifs ; ce qui est dit des textes cités vient des sources datées ci-dessus.