

Grille d'évaluation MCP : droits d'écriture

Ce qu'il faut exiger avant qu'un agent IA écrive dans vos systèmes

Demandez à chaque fournisseur de montrer chaque comportement en direct. Notez **D** démontré, **S** présenté sur une diapositive seulement, **A** absent. Considérez comme absent ce qu'il ne sait pas démontrer.

Exigence	La question à poser	Une bonne réponse	À tester en démo	Fournisseur A	Fournisseur B	Fournisseur C
1. Validation exacte	Qui valide, en voyant quoi ?	Une personne, sur les paramètres définitifs, hors de portée de l'agent. Toute modification l'annule.	Validez une écriture, modifiez un paramètre, soumettez à nouveau. L'appel doit être refusé.			
2. Classement des écritures	Qui décide qu'un outil écrit ?	Votre manifeste versionné, dans la passerelle. Outil inconnu, outil refusé.	Marquez un outil d'écriture en lecture seule. Il doit toujours exiger une validation.			
3. Exécution unique	Que fait un réessai ?	Une validation, une clé d'idempotence, au plus une exécution.	Envoyez deux fois le même appel validé, puis une fois avec un paramètre modifié, puis faites representer l'action par l'agent. Résultat attendu : une exécution et deux refus.			
4. Journal probant	Que prouve le journal ?	Des entrées en append-only, dans un stockage que personne ne peut réécrire.	Rejouez une écriture validée depuis le journal, puis demandez qui peut supprimer cet enregistrement.			
5. Droits de l'utilisateur	Quels droits porte l'appel ?	Ceux de l'utilisateur, par token exchange. Aucun compte d'agent aux droits larges.	Faites demander un remboursement à l'agent par quelqu'un qui n'a pas ce droit. Le système cible doit refuser.			
6. Comportement en panne	Et si la validation tombe ?	Les écritures s'arrêtent ; celles en attente ont un responsable.	Arrêtez le service de validation et demandez à l'agent d'écrire. Rien ne doit s'exécuter.			

Source : Rami Osman, architecte IA et cloud, Paris. « Ce qu'il faut exiger d'un serveur MCP avant qu'un agent écrive dans vos systèmes », 5 octobre 2026. <https://osmanrami.fr/fr/writing/mcp-server-security-write-access/>
Licence : CC BY 4.0. Vous pouvez partager et adapter cette grille en citant l'auteur et en renvoyant à la page source.