Construire un protocole de comparaison d’outils

JVK Formation · N7-13 · 12 novembre 2026 · Support apprenant

Objectif

Définir des tâches, des données et une grille commune avant de comparer les résultats.

Prérequis et environnement

Notions IA du N5 et contrôles N6-18/N6-19.

Cible : applications de bureau Microsoft 365 ou Office 2024 sous Windows pour les manipulations avancées. Relever la version utilisée. Les variantes web peuvent différer. Aucun script, macro, API ou échange automatique entre applications n’est demandé.

Cas à réaliser

Deux outils fictifs à comparer sur des productions préparées. Protocole et critères avant examen des résultats.

Préparer le travail

Données et situations fictives. Conserver les sources et travailler sur une copie. Les éventuels essais de partage sont simulés : aucun compte, contact ou permission réelle ne doit être modifié. Les durées et observations humaines restent à relever en situation.

Définir ce que la comparaison doit mesurer

Comparer des outils demande de préciser les tâches et le contexte avant de lire leurs réponses. Le dossier contient deux jeux de réponses entièrement simulées pour la formation, pas une mesure d’applications réelles. Les trois tâches portent sur un résumé fidèle, un calcul et une reformulation. La préférence de style ne doit pas masquer une erreur de fait. Écrivez les critères et la règle de décision avant de consulter les productions.

Assurer des conditions équivalentes

Les deux outils fictifs reçoivent les mêmes sources, les mêmes consignes et les mêmes contraintes de sortie. Un changement de consigne entre A et B rendrait la comparaison ambiguë. Conservez la version de chaque demande et de chaque réponse. L’ordre de lecture peut influencer votre préférence ; un examen anonyme peut aider. La conclusion reste liée à ces trois tâches, sans déduire une performance universelle d’un petit ensemble préparé.

Séparer exactitude, complétude et forme

La grille attribue deux points pour un résultat fidèle et complet, un point pour un résultat utilisable après une correction mineure et zéro pour une erreur décisive. Le critère critique est la fidélité aux faits et aux calculs. La fluidité peut être commentée séparément mais ne compense pas une confirmation inventée. Justifiez chaque note par un passage et par la source. Une note sans explication serait difficile à reproduire.

Expliquer les résultats et leurs limites

Dans le cas, A échoue sur la réserve d’accès mais réussit le calcul et la reformulation. B réussit le résumé mais échoue sur le calcul et ajoute une certification. Avec la règle fournie, A obtient quatre points sur six, B deux sur six. Ce résultat décrit les réponses simulées fournies, pas une supériorité générale d’un service. Proposez les tâches complémentaires nécessaires avant une décision d’usage réelle.

Exemple distinct

Comparer une réponse très courte à une réponse détaillée ne renseigne pas leur exactitude si les consignes de longueur différaient. Les conditions font partie de la preuve de comparaison.

Préparer

Rédiger le protocole, les critères et la règle 0/1/2 avant d’examiner les réponses.

Appliquer

Noter les six réponses avec la même grille. Citer le passage et le fait qui justifient chaque résultat.

Comparer

Produire une synthèse par tâche et par outil fictif. Garder les erreurs critiques visibles.

Limiter

Écrire la conclusion limitée à ces trois tâches et proposer un second jeu de tests sans prétendre l’avoir exécuté.

Pièges et reprise

Ne modifiez pas le barème pour favoriser une réponse. Ne convertissez pas trois tâches en pourcentage de fiabilité générale. Les exemples n’imposent aucun abonnement ni compte.

Critères de réussite

Protocole antérieur aux notes.

Six notes justifiées.

Totaux 4/6 et 2/6.

Portée simulée et limites explicites.

Ne pas confondre préférence de style et exactitude ; appliquer les mêmes critères aux deux outils.

Avant de remettre

Rouvrir les fichiers, vérifier les nombres et les destinations utiles, comparer la version modifiable et son PDF. Le journal distingue résultat attendu, observation réellement faite et point non testé. Un fichier ou une archive présent ne prouve pas à lui seul que l’exercice ou la transmission a réussi.

Sources et portée

CNIL — IA générative et vie privée : https://www.cnil.fr/sites/default/files/2026-05/affiche_ia_generative_et_vie_privee_fr.pdf — consultation du 12 novembre 2026.