Dimitri Jorand.

ENTREPRENDRE. TRANSMETTRE. FAIRE AVANCER.

GUIDE + TRAME DE TRAVAIL

INTELLIGENCE ARTIFICIELLE

Comparer deux réponses à l’aveugle : fiche de lecture et dossier corrigé

Un dossier fictif, deux réponses à examiner et une fiche individuelle pour argumenter votre choix. Repérez les faits, les omissions et les incertitudes avant de discuter des préférences. Le corrigé et le modèle rempli se consultent après votre première décision.

AUTEUR

Dimitri Jorand

CONTENU MIS À JOUR

23.10.2026

COLLECTION

Outils de terrain

PRISE EN MAIN

Un outil pour agir.

À QUI S'ADRESSE CE GUIDE ?

Responsables de projet, dirigeants et équipes préparant une comparaison humaine de deux productions textuelles sur une tâche documentaire limitée.

MODE D'EMPLOI

Lisez la méthode, la pièce F-07 et les deux réponses sans ouvrir le corrigé. Consultez les critères puis renseignez les champs 1 à 6 individuellement ; dans le PDF, rejoignez la trame par le sommaire. En groupe, faites commencer une partie des lecteurs par A, l'autre par B, et notez l'ordre. Comparez ensuite les observations, conservez les avis initiaux et remplissez les champs 7 et 8. Ouvrez seulement alors la section « Corrigé : à lire après votre décision ». Le guide contient toutes les pièces ; le PDF statique sert à la lecture et à l'écriture sur papier. Les exports des réponses conservent vos notes et leurs repères, sans les confondre avec le corrigé. Aucun texte n'est envoyé et aucun système n'est testé automatiquement.

VOTRE PARCOURS DANS LE DOCUMENT

01 Comprendre la méthode 3
Repères, explications et exemples pour préparer vos réponses.

02 Passer à votre situation 10
Une trame à imprimer et à compléter, séparée du guide.

03 Retrouver les références 13
Portée, limites, sources et informations de version.

IMPRIMER SEULEMENT CE QUI VOUS SERT

Pour travailler sur papier : pages 10 à 12. Ce PDF est conçu pour l'écriture à la main. Pour saisir et exporter vos réponses, utilisez la version en ligne.

Ouvrir la ressource en ligne

01 / LE GUIDE

Comprendre avant de remplir.

Ce que vous allez produire

Le livrable est une fiche de lecture argumentée, avec un avis individuel conservé et une conclusion limitée au dossier. Elle permet de comprendre ce qui a fait choisir, refuser ou suspendre une réponse. Elle ne désigne pas le meilleur outil du marché.

Tout le matériel de l'exercice figure ci-dessous : consigne, note source F-07, réponses A et B, critères et corrigé. Les deux réponses sont des compositions éditoriales fictives ; elles n'ont pas été obtenues lors d'un test comparatif de modèles. Les lecteurs du modèle rempli sont inventés, eux aussi.

La définition de critères distincts s'appuie sur un repère de recherche en évaluation humaine du texte. [1] Les huit champs, le dossier et les règles de décision sont notre proposition de travail. Ils ne reproduisent pas un instrument de mesure validé.

Préparer une lecture avant de discuter

Travail individuel. Lisez la consigne et la référence, puis les deux réponses. Remplissez les champs 1 à 6 avant d'ouvrir le corrigé. Écrivez les passages qui fondent votre avis. Si vous connaissez déjà les réponses ou leur corrigé, dites-le : vous pouvez travailler la méthode, mais votre lecture n'est plus une première comparaison.

Travail en groupe. Une personne prépare les copies et conserve l'éventuelle correspondance avec les provenances. Ici, les identifiants A et B restent les mêmes pour tous. Certains commencent par A, d'autres par B ; chacun note son ordre et lit les deux. Ne rebaptisez pas les textes entre lecteurs, au risque de discuter de productions différentes sous la même lettre.

Masquer l'identité du système et varier la présentation figurent dans les recommandations de van Miltenburg et ses coauteurs. [2] Dans cet atelier, ce sont des précautions de méthode, sans promesse d'une lecture exempte de biais. Les pièces, les critères et le temps d'examen doivent rester accessibles à chacun.

Conservez les réponses intégralement. La personne qui prépare le dossier ne corrige ni une erreur, ni une faute, ni une formulation embarrassante. Elle n'annonce pas quelle réponse elle souhaite voir retenue. Après l'avis individuel, les échanges peuvent porter sur les passages et leurs conséquences.

Pièce fictive F-07 : la référence de l’exercice

Consigne commune. Rédiger une courte note d’information destinée aux personnes intéressées par l’atelier, à partir de F-07 uniquement. Elle doit permettre de retrouver la date, le lieu, les modalités d’inscription, le matériel utile et les conditions d’accès aux supports. Conserver les éléments encore en attente. Ne confirmer aucune inscription à la place de la coordination et ne promettre aucune date absente. Aucun envoi n’est effectué dans l’exercice.

Note logistique F-07, version 2 du 20 octobre 2026. Document entièrement fictif. Les lignes ci-dessous constituent toute la référence ; aucune information extérieure n’est nécessaire pour l’exercice.

L1. L’atelier est prévu le 27 novembre 2026, de 9 h 30 à 12 h 30, en présentiel aux Aulnes, salle R2. Aucun accès à distance n’est prévu.

L2. La capacité maximale est de 12 personnes. À la date de la note, 10 inscriptions sont confirmées et 2 demandes restent à vérifier. Déposer une demande ne vaut pas confirmation d’une place.

L3. Les demandes sont recevables jusqu’au 20 novembre à 18 h. La coordination vérifie les demandes et adresse les confirmations individuelles. Les questions d’inscription lui sont transmises.

L4. Un ordinateur portable est facultatif. Les exercices se font en binôme et aucune installation n’est demandée avant l’atelier.

L5. L’atelier ne sera pas enregistré. Des supports PDF seront transmis après l’atelier, après validation par l’intervenant. La date d’envoi n’est pas fixée.

L6. Pour cette note destinée aux participants potentiels, les conditions d’inscription, le caractère facultatif du portable, l’absence d’enregistrement et l’absence de date d’envoi fixée doivent rester explicites.

Deux réponses à examiner avant le corrigé

Réponse A. Texte éditorial fictif, conservé tel quel pour la lecture.

« Nous vous retrouverons le 27 novembre 2026, de 9 h 30 à 12 h 30, aux Aulnes, salle R2, pour un atelier en présentiel. Les douze places sont désormais attribuées. Pour toute question relative à votre inscription, contactez la coordination ; les demandes restent recevables jusqu’au 20 novembre à 18 h et font l’objet d’une confirmation individuelle.

Pour profiter pleinement des exercices en binôme, munissez-vous d’un ordinateur portable. Aucune installation préalable n’est nécessaire. L’atelier ne sera pas enregistré. Après validation par l’intervenant, les supports PDF vous seront transmis dès le lendemain de l’atelier. »

Réponse B. Texte éditorial fictif, conservé tel quel pour la lecture.

« Atelier : 27 novembre 2026, 9 h 30 à 12 h 30. Lieu : les Aulnes, salle R2. Présentiel uniquement, aucun accès à distance.

Douze personnes au maximum. Dix inscriptions sont confirmées ; deux demandes restent à vérifier. Déposez votre demande avant la clôture du 20 novembre à 18 h : la demande seule ne confirme pas la place. La coordination vérifie les demandes et envoie les confirmations individuelles. Adressez-lui vos questions d'inscription.

Portable facultatif. Exercices en binôme ; aucune installation avant l'atelier. Pas d'enregistrement. Les supports PDF seront transmis après l'atelier et après validation par l'intervenant ; la date d'envoi n'est pas fixée. »

Critères communs et fiche de lecture

Pour chaque critère et chaque réponse, utilisez « respecté », « écart » ou « à clarifier », accompagné du passage précis. Le tableau définit le contrôle ; il ne donne pas les résultats.

Critère	Élément à examiner	Règle pour cet exercice
Fidélité	Date, durée, lieu, capacité et statut des inscriptions	Un fait ajouté ou modifié bloque l'usage en l'état.
Couverture	Modalités d'inscription, matériel, enregistrement et supports	Les conditions explicitement requises par L6 doivent apparaître.
Incertitude	Demandes à vérifier et date d'envoi non fixée	Aucune attente ne devient une confirmation ou une promesse.
Clarté	Compréhension du prochain geste et des conditions	Une préférence de forme ne compense pas un écart de fond.

Décisions possibles. Retenir A ; retenir B ; constater une égalité entre deux réponses acceptables ; refuser les deux ; suspendre la conclusion si une pièce ou un critère manque. « Retenir » signifie seulement que le texte satisfait ici les conditions de lecture définies. Cela ne réalise ni validation opérationnelle ni envoi.

Champ individuel	Trace attendue
1. Tâche et référence	Consigne, destinataire et version de F-07
2. Exigences	Critères et écarts bloquants avant comparaison

Champ individuel	Trace attendue
3. Présentation	Ordre lu et connaissance éventuelle de la provenance
4. A	Constats par critère, passages et effet
5. B	Constats par critère, passages et effet
6. Avis initial	Issue et motif avant toute discussion
7. Discussion	Désaccord, clarification et maintien ou révision de l’avis
8. Conclusion	Choix limité, suite et absence de généralisation

Arrêtez votre lecture ici avant le corrigé. Dans le PDF, rejoignez la trame à compléter par le sommaire et renseignez d’abord les champs 1 à 6. En groupe, discutez ensuite les observations et complétez 7 et 8. Sur le site, revenez à la trame en début de page. La section suivante révèle les réponses attendues : ne la consultez qu’après avoir conservé votre décision.

Corrigé : à lire après votre décision

Corrigé éditorial de l'exercice fictif. Il ne rapporte ni une performance de modèle ni une expérience conduite avec des participants. Chaque constat ci-dessous se vérifie dans F-07 et dans les deux textes.

Point contrôlé	Réponse A	Réponse B
Date, horaires et lieu	Fidèles à L1 ; présentiel indiqué.	Fidèles à L1 ; absence de distance explicite.
Capacité et inscriptions	« Douze places attribuées » contredit L2 : seules dix inscriptions sont confirmées.	Douze places maximum, dix confirmations et deux demandes à vérifier sont distinguées.
Dépôt et confirmation	Clôture et confirmation individuelle figurent, mais ne réparent pas l'affirmation précédente.	Clôture, vérification et absence de confirmation automatique correspondent à L2-L3.
Matériel	« Munissez-vous » ne conserve pas le caractère facultatif explicitement requis par L4-L6.	Le portable reste facultatif ; binôme et absence d'installation sont conservés.
Enregistrement	L'absence d'enregistrement correspond à L5.	L'absence d'enregistrement correspond à L5.
Supports et date	Validation conservée, mais « dès le lendemain » ajoute un engagement absent de L5.	Après atelier et validation ; date non fixée conservée.

Conclusion du corrigé. B est retenue dans le périmètre de cet exercice. A est refusée en l'état : le statut des inscriptions est modifié, la condition sur le matériel n'est pas conservée et un délai d'envoi est promis sans fondement. Les qualités de formulation que l'on peut reconnaître à A ne résolvent pas ces écarts.

Nuance de lecture. Un lecteur pourrait entendre « munissez-vous » comme un conseil plutôt que comme une obligation formelle. Ce désaccord n'oblige pas à trancher une intention supposée : la consigne L6 exige que le caractère facultatif reste explicite, ce que A ne fait pas. Cette justification est plus précise qu'un jugement sur le ton.

Une correction de A est possible, mais elle devient une nouvelle version à relire. Ne corrigez pas le texte puis ne comptez pas sa version corrigée comme si c'était la sortie initiale. Dans un véritable essai, la charge de reprise reste également à examiner.

EXEMPLE FICTIF / POUR COMPRENDRE

Exemple fictif entièrement rempli : de l’avis individuel à la conclusion

Les notes suivantes montrent comment compléter les huit rubriques. Les rubriques 1 à 3 décrivent le cadre commun ; les rubriques 4 et 5 reproduisent les constats individuels du lecteur 2. La rubrique 6 conserve les deux avis initiaux, puis les rubriques 7 et 8 leur discussion et sa conclusion. Les lecteurs et leur discussion sont inventés ; aucune évaluation réelle n’est rapportée.

1. Tâche et référence. Préparer une information destinée aux personnes intéressées par l’atelier fictif. Référence : F-07, version 2 du 20 octobre 2026, lignes L1 à L6. Décider quel texte conserve les conditions annoncées, sans envoyer de message. Aucune autre pièce n’est utilisée.

2. Exigences. Fidélité des faits ; présence des conditions requises par L6 ; maintien des attentes ; clarté. Un statut d’inscription changé, une condition essentielle absente ou une date promise sans source bloque l’usage en l’état. Pas de note globale permettant de compenser ces écarts par le style.

3. Présentation. Le lecteur 1 lit A puis B ; le lecteur 2, B puis A. Tous deux disposent de F-07 et des mêmes critères. Les textes ne sont pas corrigés. Il n’existe aucune identité de modèle cachée : ce sont deux variantes éditoriales, comme annoncé dès l’ouverture.

4. Constats du lecteur 2 sur A, avant discussion. Date, lieu, binôme et absence d’enregistrement conformes. Trois points à reprendre : « douze places attribuées » contre les dix confirmations de L2 ; portable facultatif non explicite contrairement à L6 ; lendemain promis alors que L5 ne fixe aucune date. Le texte ne peut être utilisé en l’état.

5. Constats du lecteur 2 sur B, avant discussion. Les conditions de F-07 sont conservées, y compris deux demandes à vérifier, portable facultatif et date d’envoi non fixée. Le style est plus télégraphique, mais l’action d’inscription et ses limites restent compréhensibles. Aucun écart bloquant relevé dans ce dossier.

6. Avis individuels avant discussion. Le lecteur 1 écrit d’abord « A, formulation plus accueillante » ; sa fiche ne contient alors qu’une appréciation du ton et aucun examen des écarts de fond. Le lecteur 2 écrit « B, seule réponse conservant les statuts et conditions ». Ce sont des avis pédagogiques inventés, pas des données d’essai.

7. Désaccord et clarification. Retour collectif à L2, L5 et L6. Le lecteur 1 examine à son tour les écarts, complète sa fiche en les rattachant aux lignes de F-07, puis distingue sa préférence de ton de la décision d’usage et révise son choix vers B. Son avis initial reste visible. Le désaccord sur le caractère plus agréable de A demeure, sans modifier le constat factuel.

8. Conclusion, suite et limite. B retenue pour ce seul dossier fictif. A non retenue en l'état ; toute réécriture constituerait une version distincte. Aucun envoi, aucune mesure de temps et aucun modèle évalué. Suite possible : préparer un autre cas qui fasse apparaître une omission différente, sans conclure à une fiabilité générale.

Égalité, refus des deux et adaptation à votre situation

Égalité acceptable. Si deux versions respectent toutes les conditions et que leur différence de forme n'est pas déterminante pour le destinataire, conservez une égalité. « Je n'arrive pas à vérifier » n'est pas une égalité : c'est une conclusion suspendue.

Refus des deux. Variante pédagogique : imaginez que B soit réécrite avec « ordinateur obligatoire », tandis que A reste inchangée. L4 et L6 rendent alors les deux textes impropres à l'usage prévu. Il faut les reprendre, même si l'un contient moins d'erreurs. Cette variante est un raisonnement éditorial, pas une nouvelle sortie obtenue d'un outil.

Pièce ou règle ambiguë. Si la référence elle-même est contradictoire, conservez le point ouvert et faites-la clarifier avant de prétendre départager les réponses. Si vous changez un critère, appliquez sa nouvelle version à A et à B, en gardant la trace du changement.

Pour vos cas, remplacez F-07 par un dossier autorisé et préparez les attentes avant de lire les sorties. Gardez séparément les textes originaux, la correspondance de provenance et les avis. Ne mélangez pas plusieurs versions d'une consigne sous le même identifiant.

Le NIST rappelle la limite des généralisations tirées d'évaluations étroites ou anecdotiques. [3] Une fiche remplie constitue une trace de lecture sur un cas, pas une preuve de performance sur tous les usages. La méthode associée explique comment organiser la comparaison humaine.

TRAME À COMPLÉTER / 1 SUR 3

À vous de jouer.

Comparer deux réponses à l'aveugle : fiche de lecture et dossier corrigé

Projet / situation : _____ Date : _____

01 Quelle tâche, quel destinataire et quelle version du dossier examinez-vous ?

Indiquez la référence des pièces, la consigne et l'état final attendu : brouillon à reprendre ou texte utilisable après le contrôle défini. Pour l'exercice : note F-07, version 2, information aux personnes intéressées par l'atelier fictif.

02 Quels critères doivent être satisfaits, et quels écarts empêchent l'usage prévu ?

Définissez fidélité, couverture, incertitudes et clarté avant de juger. Distinguez exigences indispensables et préférences. Ne changez pas les règles seulement pour départager les réponses ; si elles changent, réexaminez A et B.

03 Comment la provenance et l'ordre de présentation ont-ils été traités ?

Notez les identifiants, l'ordre réellement lu, ce qui a été masqué et ce qui était déjà connu. Conservez les textes. Dans cet exercice, A/B sont des compositions éditoriales sans modèle à révéler ; en situation réelle, la correspondance de provenance reste à part jusqu'à la fin.

TRAME À COMPLÉTER / 2 SUR 3

À vous de jouer.

Comparer deux réponses à l'aveugle : fiche de lecture et dossier corrigé

Projet / situation : _____ Date : _____

04 Que constatez-vous dans A, critère par critère, avec quels passages de référence ?

Pour chaque critère, écrivez respecté, écart ou à clarifier, puis un passage exact et son effet sur l'usage. Ne réécrivez pas A pour lui attribuer ensuite la qualité de votre correction. Les lignes L1 à L6 permettent de vérifier les faits.

05 Que constatez-vous dans B, critère par critère, avec quels passages de référence ?

Appliquez exactement les mêmes critères et le même niveau d'exigence. Relevez aussi ce qui reste à contrôler ; une préférence pour B n'est pas une preuve générale de fiabilité.

TRAME À COMPLÉTER / 3 SUR 3

À vous de jouer.

Comparer deux réponses à l'aveugle : fiche de lecture et dossier corrigé

Projet / situation : _____ Date : _____

06 Quelle est votre décision individuelle avant discussion, et pourquoi ?

Choisissez A, B, égalité acceptable, refus des deux ou conclusion suspendue. Distinguez préférence de style et acceptabilité pour la tâche. Enregistrez cet avis avant de lire les avis des autres ou le corrigé.

07 Quels désaccords ont été discutés, avec quelles raisons conservées ou corrigées ?

Conservez le motif initial et la pièce qui a fait évoluer le jugement. Une divergence de style peut demeurer. Un critère ambigu appelle une clarification appliquée aux deux textes, pas un vote destiné à faire disparaître l'écart.

08 Quelle conclusion retenez-vous pour ce cas, avec quelle suite et quelles limites ?

Nommez la version et le cas concernés, la correction éventuelle, la personne qui doit vérifier et la prochaine étape. N'attribuez pas le choix à un modèle, un taux de réussite ou une population de demandes qui n'ont pas été examinés.

03 / POUR ALLER PLUS LOIN

Portée, sources et version.

PORTÉE ET LIMITES

Exercice original entièrement fictif : textes A/B et avis inventés, sans test de logiciel ni modèle évalué. Aucune performance, validation scientifique, conformité ou aptitude au déploiement démontrée. Hors décisions médicales, juridiques et de sécurité. Utilisez des pièces fictives ou autorisées, sans information personnelle ni confidentielle.

Sources et méthode

[1] Van der Lee et al. : Best practices for the human evaluation of automatically generated text (2019)

PDF consulté le 22 octobre 2026, sections 4.1 et 4.4, tableau 3 : critères distincts et ordre de présentation. Notre fiche est une adaptation originale, pas un protocole expérimental validé.

https://aclanthology.org/W19-8643/

[2] Van Miltenburg et al. : Underreporting of errors in NLG output, and what to do about it (2021)

PDF consulté le 22 octobre 2026, section 4.2, Presentation et Role of the annotators : identité masquée et examen des erreurs. Aucun effet chiffré repris ; textes A/B fictifs.

https://aclanthology.org/2021.inlg-1.14/

[3] NIST AI 600-1 : Generative Artificial Intelligence Profile (juillet 2024)

Consulté le 22 octobre 2026, action MS-2.5-001, page 31 : éviter les généralisations d'évaluations étroites ou anecdotiques. Cadre volontaire américain, sans certification ni examen de conformité.

https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

Informations sur cette édition

Auteur : Dimitri Jorand.

Première publication : 23.10.2026. Contenu mis à jour : 23.10.2026. Édition graphique : 21.10.2026.

Production éditoriale assistée par IA. Cette signature ne signifie pas que Dimitri a personnellement relu chaque publication. Les cas fictifs servent à comprendre la méthode ; ils ne décrivent pas des résultats clients.

Consulter la ligne éditoriale · Signaler une correction

La version en ligne est le point de référence pour les mises à jour.

https://dimitri-jorand.fr/ressources/fiche-lecture-aveugle-reponses-ia/