ask this guy

# Qualité du français écrit par les LLM

Run `fr1` : 48 consignes, 2 génération(s) par consigne, 6 modèles (576 textes), mis à jour le 2026-11-24T14:16:56+00:00.

La méthode (corpus, contrôles, juges, statistiques) et le texte intégral des consignes figurent en fin de rapport.

## Principaux résultats

- Gemini : Note globale : meilleur DeepSeek V4.1 Flash (4,90), plus faible Qwen 3.8 27B (4,43).
- Jev : Fluidité : meilleur GPT-6 Luna (4,52), plus faible Qwen 3.8 27B (4,37).
- Aucun modèle n'a écrit de caractères d'un autre alphabet.
- Alertes de grammaire de Grammalecte pour 1 000 mots : moins nombreuses chez GPT-6 Luna (0,6), plus nombreuses chez Qwen 3.8 27B (1,7).
- Alertes d'orthographe de Grammalecte pour 1 000 mots : moins nombreuses chez GPT-6 Luna (0,3), plus nombreuses chez Qwen 3.8 27B (2,4).
- Pas de biais significatif de Gemini en faveur de Gemma 4 31B (modèle Google) par rapport à Jev (+0,09 [-0,05 ; +0,22]).

## Modèles

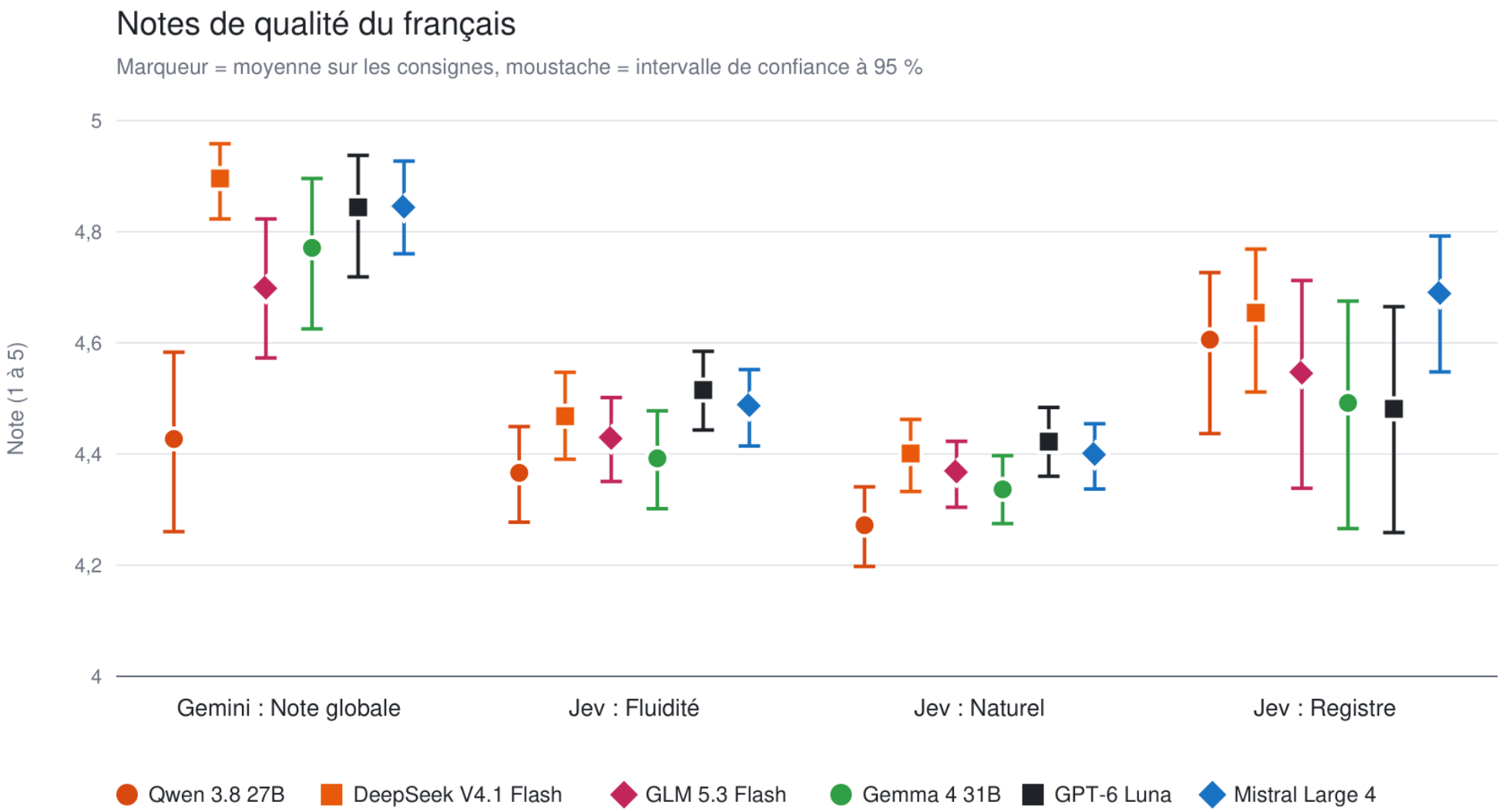
| Modèle | Éditeur | Famille | Fournisseur | ID | Raisonnement | Prix par M tokens (entrée / sortie) |
|---|---|---|---|---|---|---|
| Qwen 3.8 27B | Alibaba | Open-weight chinois | OVHcloud AI Endpoints | `Qwen3.8-27B` | `reasoning_effort=none` | 0,47 / 3,19 USD |
| DeepSeek V4.1 Flash | DeepSeek | Open-weight chinois | OVHcloud AI Endpoints | `DeepSeek-V4.1-Flash-1017` | `reasoning_effort=none` | 0,31 / 1,25 USD |
| GLM 5.3 Flash | Zhipu AI | Open-weight chinois | Melious | `glm-5.3-flash` | `reasoning_effort=low` | 0,10 / 0,40 EUR |
| Gemma 4 31B | Google | Open-weight américain | Melious | `gemma-4-31b` | `reasoning_effort=none` | 0,10 / 0,30 EUR |
| GPT-6 Luna | OpenAI | Fermé américain | OpenAI | `gpt-6-luna` | `reasoning_effort=none` | 0,10 / 0,50 USD |
| Mistral Large 4 | Mistral AI | Open-weight français | Mistral AI | `mistral-large-4` | `reasoning_effort=low` | 0,70 / 2,10 USD |

## Notes de qualité

Moyenne sur les consignes (pour chaque consigne : moyenne des générations) et intervalle de confiance à 95 % (bootstrap), sur une échelle de 1 à 5. Les notes Jev sont des scores espérés (pondérés par les probabilités) ; les notes Gemini sont entières.

| Modèle | Gemini : Correction | Gemini : Naturel | Gemini : Registre | Gemini : Typographie | Gemini : Note globale | Jev : Fluidité | Jev : Naturel | Jev : Registre |
|---|---|---|---|---|---|---|---|---|
| DeepSeek V4.1 Flash | 4,95 [4,89 ; 4,99] | 4,92 [4,85 ; 4,97] | 4,96 [4,91 ; 5,00] | 4,71 [4,59 ; 4,81] | 4,90 [4,82 ; 4,96] | 4,47 [4,39 ; 4,55] | 4,40 [4,33 ; 4,46] | 4,65 [4,51 ; 4,77] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GPT-6 Luna | 4,96 [4,91 ; 5,00] | 4,98 [4,95 ; 5,00] | 4,88 [4,72 ; 4,99] | 4,75 [4,64 ; 4,84] | 4,84 [4,72 ; 4,94] | 4,52 [4,44 ; 4,58] | 4,42 [4,36 ; 4,48] | 4,48 [4,26 ; 4,67] |
| Mistral Large 4 | 4,92 [4,85 ; 4,97] | 4,94 [4,88 ; 4,98] | 4,96 [4,91 ; 5,00] | 4,61 [4,50 ; 4,73] | 4,84 [4,76 ; 4,93] | 4,49 [4,41 ; 4,55] | 4,40 [4,34 ; 4,45] | 4,69 [4,55 ; 4,79] |
| Gemma 4 31B | 4,93 [4,85 ; 4,99] | 4,90 [4,80 ; 4,97] | 4,84 [4,65 ; 4,99] | 4,48 [4,36 ; 4,59] | 4,77 [4,62 ; 4,90] | 4,39 [4,30 ; 4,48] | 4,34 [4,27 ; 4,40] | 4,49 [4,27 ; 4,68] |
| GLM 5.3 Flash | 4,74 [4,64 ; 4,83] | 4,86 [4,79 ; 4,94] | 4,91 [4,76 ; 5,00] | 4,60 [4,51 ; 4,71] | 4,70 [4,57 ; 4,82] | 4,43 [4,35 ; 4,50] | 4,37 [4,30 ; 4,42] | 4,55 [4,34 ; 4,71] |
| Qwen 3.8 27B | 4,52 [4,38 ; 4,66] | 4,60 [4,47 ; 4,72] | 4,89 [4,74 ; 4,99] | 4,51 [4,40 ; 4,62] | 4,43 [4,26 ; 4,58] | 4,37 [4,28 ; 4,45] | 4,27 [4,20 ; 4,34] | 4,61 [4,44 ; 4,73] |



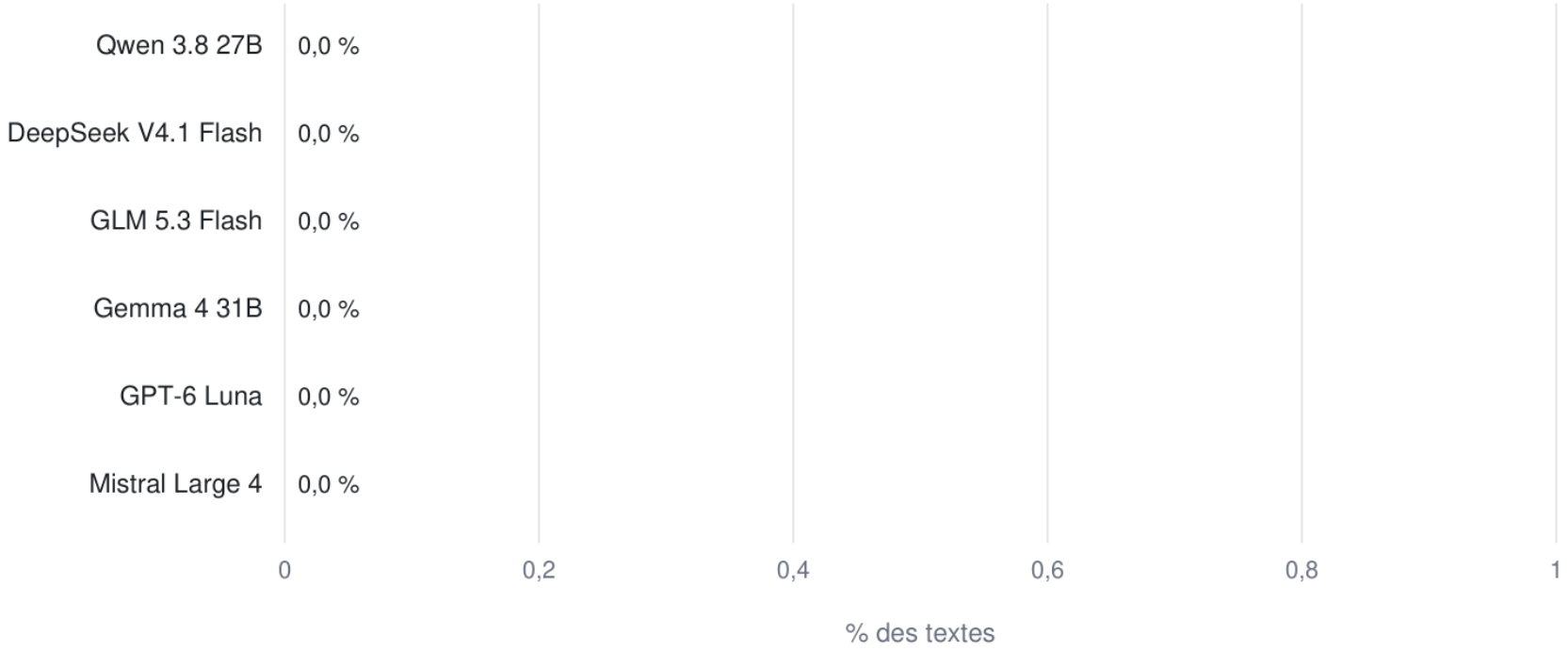
## Contrôles automatiques

Autre alphabet : textes avec au moins un caractère d'un alphabet non latin autre que le grec. Anglais : textes dont plus de 5 % des caractères sont dans des passages en anglais (lingua). Grammalecte, LanguageTool et typographie : alertes pour 1 000 mots (les mots de la consigne, comme les noms fictifs, ne comptent pas comme fautes d'orthographe). L'API publique de LanguageTool classe ses règles de grammaire françaises comme non catégorisées : elles sont dans « LT autres ».

| Modèle | Autre alphabet (textes) | Caractères CJK | Caractères cyrilliques | Anglais (textes) | Grammalecte orthographe | Grammalecte grammaire | LT orthographe | LT autres | Typographie | Longueur / cible |
|---|---|---|---|---|---|---|---|---|---|---|
| Qwen 3.8 27B | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 2,4 | 1,7 | 3,6 | 13,7 | 0,4 | 0,93 |
| DeepSeek V4.1 Flash | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 0,5 | 1,5 | 2,1 | 7,7 | 0,3 | 0,82 |
| GLM 5.3 Flash | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 1,7 | 0,6 | 3,1 | 7,0 | 0,7 | 1,11 |
| Gemma 4 31B | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 0,9 | 1,3 | 2,6 | 15,0 | 1,6 | 0,78 |
| GPT-6 Luna | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 0,3 | 0,6 | 1,8 | 7,8 | 0,0 | 0,75 |
| Mistral Large 4 | 0,0 % (0) | 0 | 0 | 0,0 % (0) | 1,0 | 1,1 | 2,3 | 8,2 | 0,1 | 0,79 |

## Textes avec des caractères d'un autre alphabet

Chinois, cyrillique, arabe... (grec exclu)




Fautes de typographie pour 1 000 mots, par règle :

| Modèle | Pas d'espace avant ; : ! ? | Guillemets droits " | Nombres au format anglais |
|---|---|---|---|
| Qwen 3.8 27B | 0,11 | 0,33 | 0,00 |
| DeepSeek V4.1 Flash | 0,07 | 0,27 | 0,00 |
| GLM 5.3 Flash | 0,00 | 0,67 | 0,00 |
| Gemma 4 31B | 0,26 | 1,32 | 0,00 |
| GPT-6 Luna | 0,00 | 0,00 | 0,00 |
| Mistral Large 4 | 0,00 | 0,14 | 0,00 |

## Retours de Gemini par catégorie

Problèmes signalés par Gemini pour 1 000 mots (majeurs / mineurs). Extraits introuvables : part des retours dont l'extrait cité n'est pas dans le texte.

| Modèle | Grammaire | Orthographe | Anglicisme | Calque | Registre | Typographie | Autre alphabet | Autre | Majeurs, total | Extraits introuvables |
|---|---|---|---|---|---|---|---|---|---|---|
| Qwen 3.8 27B | 0,7 / 1,5 | 0,3 / 0,2 | 0,1 / 0,8 | 0,0 / 0,3 | 0,2 / 0,2 | 0,0 / 3,1 | 0,0 / 0,0 | 0,3 / 1,4 | 1,7 | 4 % |
| DeepSeek V4.1 Flash | 0,2 / 0,3 | 0,0 / 0,1 | 0,0 / 0,5 | 0,0 / 0,1 | 0,0 / 0,2 | 0,0 / 2,4 | 0,0 / 0,0 | 0,0 / 0,2 | 0,2 | 0 % |
| GLM 5.3 Flash | 0,2 / 1,0 | 0,4 / 0,3 | 0,0 / 0,4 | 0,0 / 0,0 | 0,2 / 0,1 | 0,0 / 2,4 | 0,0 / 0,0 | 0,1 / 0,4 | 1,0 | 6 % |
| Gemma 4 31B | 0,1 / 0,4 | 0,1 / 0,1 | 0,1 / 0,6 | 0,0 / 0,1 | 0,8 / 0,3 | 0,0 / 4,9 | 0,1 / 0,0 | 0,1 / 0,1 | 1,2 | 4 % |
| GPT-6 Luna | 0,2 / 0,1 | 0,0 / 0,1 | 0,0 / 0,0 | 0,0 / 0,0 | 0,9 / 0,2 | 0,0 / 2,2 | 0,0 / 0,0 | 0,0 / 0,1 | 1,2 | 9 % |
| Mistral Large 4 | 0,3 / 0,5 | 0,0 / 0,1 | 0,0 / 0,5 | 0,0 / 0,0 | 0,0 / 0,3 | 0,0 / 3,3 | 0,0 / 0,0 | 0,0 / 0,3 | 0,3 | 6 % |

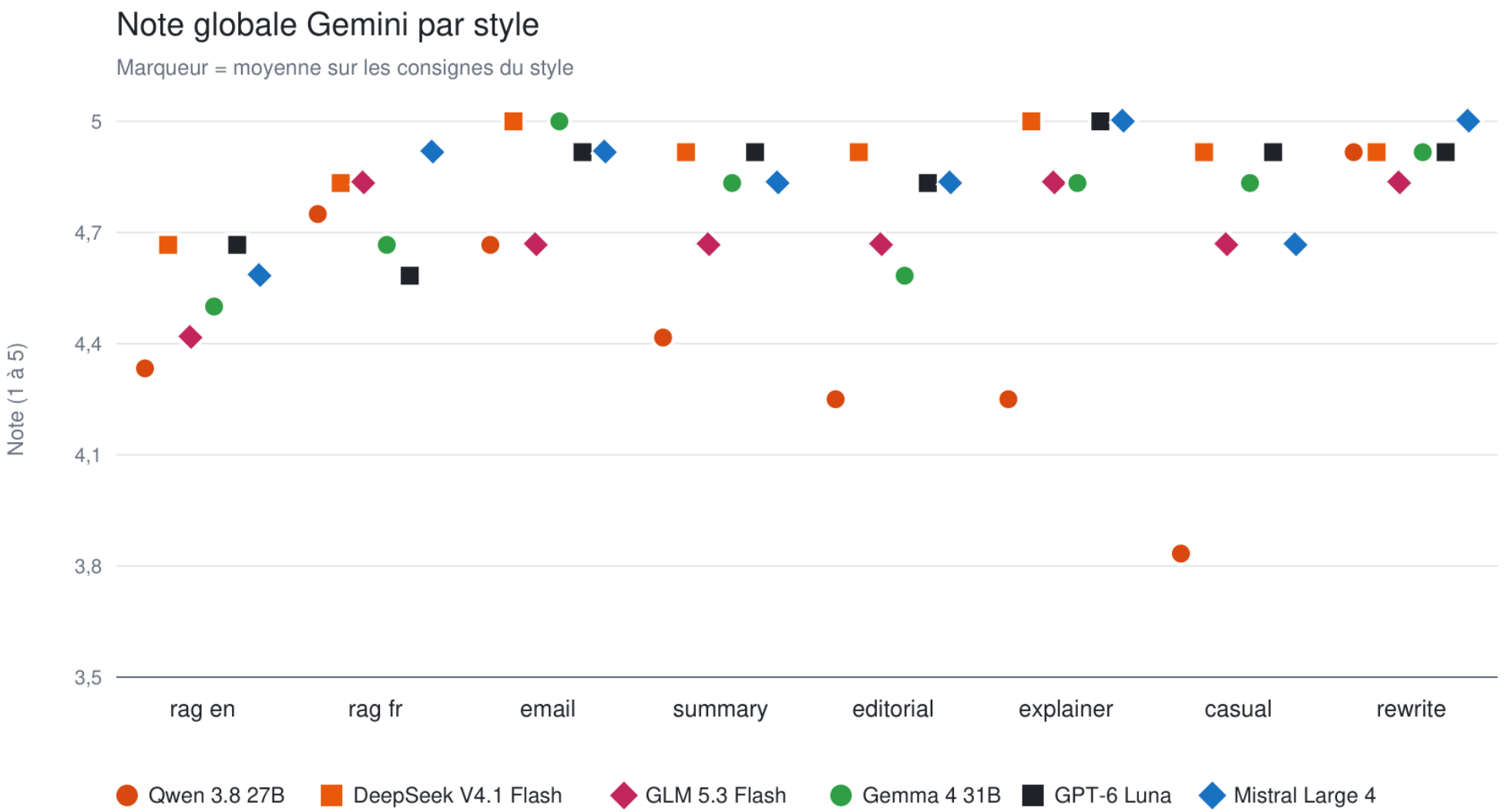
## Par style

Note globale Gemini (moyenne sur les consignes du style).

| Style | Qwen 3.8 27B | DeepSeek V4.1 Flash | GLM 5.3 Flash | Gemma 4 31B | GPT-6 Luna | Mistral Large 4 |
|---|---|---|---|---|---|---|
| Réponse à partir d'extraits en anglais | 4,33 | 4,67 | 4,42 | 4,50 | 4,67 | 4,58 |
| Réponse à partir d'extraits en français | 4,75 | 4,83 | 4,83 | 4,67 | 4,58 | 4,92 |
| E-mail ou courrier professionnel | 4,67 | 5,00 | 4,67 | 5,00 | 4,92 | 4,92 |
| Synthèse d'un document long | 4,42 | 4,92 | 4,67 | 4,83 | 4,92 | 4,83 |
| Contenu éditorial | 4,25 | 4,92 | 4,67 | 4,58 | 4,83 | 4,83 |
| Explication pédagogique | 4,25 | 5,00 | 4,83 | 4,83 | 5,00 | 5,00 |
| Échange informel | 3,83 | 4,92 | 4,67 | 4,83 | 4,92 | 4,67 |
| Reformulation | 4,92 | 4,92 | 4,83 | 4,92 | 4,92 | 5,00 |

Fluidité Jev (moyenne sur les consignes du style).

| Style | Qwen 3.8 27B | DeepSeek V4.1 Flash | GLM 5.3 Flash | Gemma 4 31B | GPT-6 Luna | Mistral Large 4 |
|---|---|---|---|---|---|---|
| Réponse à partir d'extraits en anglais | 3,91 | 4,07 | 4,04 | 4,08 | 4,26 | 4,19 |
| Réponse à partir d'extraits en français | 4,47 | 4,47 | 4,51 | 4,51 | 4,42 | 4,47 |
| E-mail ou courrier professionnel | 4,62 | 4,66 | 4,62 | 4,67 | 4,68 | 4,71 |
| Synthèse d'un document long | 4,25 | 4,33 | 4,28 | 4,15 | 4,37 | 4,22 |
| Contenu éditorial | 4,44 | 4,71 | 4,52 | 4,41 | 4,68 | 4,63 |
| Explication pédagogique | 4,36 | 4,62 | 4,60 | 4,50 | 4,75 | 4,62 |
| Échange informel | 4,23 | 4,28 | 4,29 | 4,14 | 4,38 | 4,44 |
| Reformulation | 4,66 | 4,61 | 4,57 | 4,68 | 4,58 | 4,61 |




## Comparaisons deux à deux

Différence appariée (premier modèle moins second) sur les consignes traitées par les deux, avec son intervalle à 95 % ; en gras quand elle est significative.

| Comparaison | Gemini, note globale | Jev, fluidité | Jev, naturel |
|---|---|---|---|
| Qwen 3.8 27B - DeepSeek V4.1 Flash | **-0,47 [-0,65 ; -0,29]** | **-0,10 [-0,16 ; -0,04]** | **-0,13 [-0,18 ; -0,08]** |

| | | | |
|---|---|---|---|
| Qwen 3.8 27B - GLM 5.3 Flash | **-0,27 [-0,46 ; -0,07]** | -0,06 [-0,12 ; +0,00] | **-0,10 [-0,15 ; -0,05]** |
| Qwen 3.8 27B - Gemma 4 31B | **-0,34 [-0,54 ; -0,12]** | -0,03 [-0,09 ; +0,04] | **-0,06 [-0,12 ; -0,01]** |
| Qwen 3.8 27B - GPT-6 Luna | **-0,42 [-0,60 ; -0,22]** | **-0,15 [-0,23 ; -0,08]** | **-0,15 [-0,21 ; -0,09]** |
| Qwen 3.8 27B - Mistral Large 4 | **-0,42 [-0,56 ; -0,27]** | **-0,12 [-0,18 ; -0,06]** | **-0,13 [-0,18 ; -0,08]** |
| DeepSeek V4.1 Flash - GLM 5.3 Flash | **+0,20 [+0,07 ; +0,34]** | +0,04 [-0,01 ; +0,09] | +0,03 [-0,01 ; +0,07] |
| DeepSeek V4.1 Flash - Gemma 4 31B | +0,12 [-0,03 ; +0,29] | **+0,08 [+0,01 ; +0,15]** | **+0,06 [+0,01 ; +0,12]** |
| DeepSeek V4.1 Flash - GPT-6 Luna | +0,05 [-0,04 ; +0,18] | **-0,05 [-0,09 ; -0,00]** | -0,02 [-0,06 ; +0,02] |
| DeepSeek V4.1 Flash - Mistral Large 4 | +0,05 [-0,03 ; +0,14] | -0,02 [-0,07 ; +0,03] | +0,00 [-0,04 ; +0,04] |
| GLM 5.3 Flash - Gemma 4 31B | -0,07 [-0,22 ; +0,07] | +0,04 [-0,02 ; +0,10] | +0,03 [-0,01 ; +0,08] |
| GLM 5.3 Flash - GPT-6 Luna | **-0,15 [-0,27 ; -0,01]** | **-0,09 [-0,15 ; -0,03]** | **-0,05 [-0,11 ; -0,01]** |
| GLM 5.3 Flash - Mistral Large 4 | **-0,15 [-0,28 ; -0,01]** | **-0,06 [-0,12 ; -0,01]** | -0,03 [-0,07 ; +0,01] |
| Gemma 4 31B - GPT-6 Luna | -0,07 [-0,20 ; +0,05] | **-0,12 [-0,21 ; -0,05]** | **-0,09 [-0,16 ; -0,03]** |
| Gemma 4 31B - Mistral Large 4 | -0,07 [-0,23 ; +0,07] | **-0,09 [-0,17 ; -0,03]** | **-0,06 [-0,12 ; -0,02]** |
| GPT-6 Luna - Mistral Large 4 | +0,00 [-0,12 ; +0,10] | +0,03 [-0,02 ; +0,07] | +0,02 [-0,02 ; +0,06] |

## Juges : accord et biais de famille

Accord : pour chaque consigne, corrélation de Spearman entre les notes globales Gemini des modèles et leurs notes Jev, moyenne sur les consignes. Écart : note globale Gemini moins fluidité Jev, moyenne par modèle. Gemini est un modèle Google, comme Gemma : un écart de Gemma nettement au-dessus de celui des autres modèles suggérerait un biais de famille.

- Gemini note globale vs Jev : Fluidité : 0,35 [0,21 ; 0,47]
- Gemini note globale vs Jev : Naturel : 0,35 [0,21 ; 0,47]
- Gemini note globale vs Jev : Registre : 0,23 [0,09 ; 0,37]

| Modèle | Écart Gemini - Jev |
|---|---|
| Qwen 3.8 27B | +0,06 [-0,08 ; +0,19] |
| DeepSeek V4.1 Flash | **+0,43 [+0,33 ; +0,52]** |
| GLM 5.3 Flash | **+0,27 [+0,13 ; +0,40]** |
| Gemma 4 31B | **+0,38 [+0,23 ; +0,52]** |
| GPT-6 Luna | **+0,33 [+0,20 ; +0,44]** |
| Mistral Large 4 | **+0,36 [+0,26 ; +0,45]** |

Écart de Gemma 4 31B moins l'écart moyen des autres modèles : +0,09 [-0,05 ; +0,22] (pas de biais significatif).

## Classement humain

Pas encore importé : ouvrir `human_review.html`, classer les réponses, exporter le JSON et relancer le script avec `--human-rankings`.

## Variabilité entre générations

Différence absolue moyenne entre deux générations d'une même consigne (toutes les paires de générations).

| Modèle | Gemini, note globale | Jev, fluidité |
|---|---|---|
| Qwen 3.8 27B | 0,44 | 0,18 |
| DeepSeek V4.1 Flash | 0,12 | 0,13 |
| GLM 5.3 Flash | 0,31 | 0,21 |
| Gemma 4 31B | 0,08 | 0,13 |
| GPT-6 Luna | 0,15 | 0,09 |
| Mistral Large 4 | 0,19 | 0,17 |

## Exemples

### Qwen 3.8 27B

- `casual_01#2` Autre (majeur) : « La cuisine de rue (croques, lyonnaise, quiches) » → « La cuisine locale ou les petits restos sur le pouce (bouchons, croques, quiches) ». L'énumération est bancale : « lyonnaise » est un adjectif employé seul sans nom, et le Vieux-Lyon est réputé pour ses bouchons plutôt que pour de la « cuisine de rue ».
- `casual_03#1` Autre (majeur) : « Le gros projet est officiellement dans les tuyaux » → « Le gros projet est officiellement bouclé ». « Être dans les tuyaux » signifie que le projet est en cours de préparation ou d'acheminement, et non qu'il est terminé comme demandé dans la consigne.
- `casual_03#1` Grammaire (majeur) : « voir qui est le plus adroite de l'équipe » → « voir qui est le plus adroit de l'équipe ». Le pronom « qui » a une valeur indéfinie ou masculine neutre ici, le superlatif s'accorde au masculin (« le plus adroit »).
- `casual_05#1` Registre (majeur) : « visez plus court en gardant uniquement les mots-clés sur vos slides, et essayez » → « vise plus court en gardant uniquement les mots-clés sur tes slides, et essaie ». Le texte passe brutalement du tutoiement au vouvoiement au milieu du paragraphe.

### DeepSeek V4.1 Flash

- `rag_en_05#1` Grammaire (majeur) : « tous les 8 000 heures de fonctionnement » → « toutes les 8 000 heures de fonctionnement ». Le mot « heure » est féminin, l'adjectif indéfini doit donc s'accorder au féminin pluriel (« toutes »).
- `rag_en_05#2` Grammaire (majeur) : « tous les 8 000 heures de service » → « toutes les 8 000 heures de service ». Le mot « heure » est féminin, l'adjectif indéfini doit donc s'accorder au féminin pluriel (« toutes »).

- `summary_06#2` Grammaire (majeur) : « aucune représailles contre un lanceur d'alerte » → « aucunes représailles contre un lanceur d'alerte (ou : aucune mesure de représailles contre un lanceur d'alerte) ». Le mot « représailles » est un nom exclusivement féminin pluriel ; il faut donc accorder le déterminant au pluriel (« aucunes ») ou employer « aucune mesure de représailles ».
- `casual_01#1` Registre (mineur) : « mais évitez les pièges à touristes, éloignez-vous un peu » → « mais évite les pièges à touristes, éloigne-toi un peu (ou : mais évitez les pièges à touristes à deux) ». Le passage soudain du tutoiement (« Va », « Réserve ») au pluriel (« évitez », « éloignez-vous ») crée une rupture inutile, même si l'ami est avec sa copine.

**GLM 5.3 Flash**

- `casual_01#1` Autre (majeur) : « Bonne décision ? » → « Bon séjour ! ». La formule « Bonne décision ? » en clôture est totalement inappropriée et semble être une mauvaise traduction ou une hallucination du modèle.
- `editorial_01#2` Orthographe (majeur) : « sur de petites surfaces closées, très éloignées » → « sur de petites surfaces closes, très éloignées ». Le participe passé ou adjectif du verbe clore est « clos / close », le mot « closé » n'existe pas en français.
- `editorial_01#2` Grammaire (majeur) : « Aérer quand il fait dehors peut sembler » → « Aérer quand il fait froid dehors peut sembler ». L'omission de l'adjectif « froid » rend la proposition incomplète et agrammaticale.
- `editorial_01#2` Autre (majeur) : « La facture de chauffage en pousse à peine, quand l'air » → « La facture de chauffage en pâtit à peine, quand l'air ». L'expression « en pousse à peine » est un contresens ou une coquille pour « n'en pâtit à peine » ou « n'en souffre à peine ».

**Gemma 4 31B**

- `editorial_01#1` Autre alphabet (majeur) : « une approche cohérente et sederhana. » → « une approche cohérente et simple. ». Le mot indonésien « sederhana » s'est manifestement glissé par erreur à la place de « simple ».
- `editorial_05#1` Anglicisme (majeur) : « Dans l'imaginaire collectif, reducing le temps de travail » → « réduire le temps de travail ». Le mot anglais « reducing » a été laissé par erreur au lieu de l'infinitif français « réduire ».
- `editorial_05#2` Autre (majeur) : « On a longtemps chromosomalisé la semaine de quatre jours » → « On a longtemps considéré la semaine de quatre jours ». Le verbe « chromosomaliser » n'existe pas en français et constitue un néologisme absurde dans ce contexte.
- `explainer_02#1` Orthographe (majeur) : « Pourquoi consomre-t-elle moins qu'un chauffage électrique ? » → « Pourquoi consomme-t-elle moins qu'un chauffage électrique ? ». Coquille manifeste avec la lettre « r » à la place du « m » dans le verbe « consommer ».

**GPT-6 Luna**

- `email_06#1` Grammaire (majeur) : « nos échanges nous ont laissé une excellente impression » → « nos échanges nous ont laissée une excellente impression ». Le participe passé « laissé » s'accorde avec le complément d'objet direct « impression » (féminin singulier) placé après le verbe attributif ou ici avec le sens de laisser une impression (en réalité, « laisser » suivi d'un COD postposé ne s'accorde pas : « laisser [quoi ?] une excellente impression [à qui ?] à nous »... Attendez : « laisser » a pour COD « une excellente impression », donc le participe passé reste invariable). Non, « nous » est COI (« ont laissé à nous ») et le COD est « une excellente impression » placé après. Donc « laissé » est invariable et correct !

- `rag_en_04#2` Registre (majeur) : « Tu peux installer cette application sur ton téléphone personnel : » → « Vous pouvez installer cette application sur votre téléphone personnel : ». Emploi indu du tutoiement à l'adresse d'un collaborateur au lieu du vouvoiement demandé.
- `rag_en_04#2` Registre (majeur) : « Si tu ne souhaites pas utiliser ton téléphone, demande » → « Si vous ne souhaitez pas utiliser votre téléphone, demandez ». Le vouvoiement professionnel doit être appliqué de manière cohérente.
- `rag_en_04#2` Registre (majeur) : « Pour ton mot de passe, il n'y a pas de changement périodique obligatoire. Change-le seulement » → « Pour votre mot de passe, il n'y a pas de changement périodique obligatoire. Changez-le seulement ». Maintien erroné du tutoiement au lieu du vouvoiement.

**Mistral Large 4**

- `editorial_04#1` Grammaire (majeur) : « Engagé dans la transition énergétique du secteur industriel, l'entreprise développe » → « Engagée dans la transition énergétique du secteur industriel, l'entreprise développe ». Le participe apposé « Engagé » doit s'accorder au féminin avec le sujet « l'entreprise ».
- `editorial_05#1` Grammaire (majeur) : « Parce qu'un salarié épanoui usiner mieux qu'un salarié épuisé. » → « Parce qu'un salarié épanoui usine mieux qu'un salarié épuisé. ». Le verbe doit être conjugué au présent de l'indicatif (usine) et non à l'infinitif.
- `casual_01#2` Grammaire (mineur) : « Vas plutôt » → « Va plutôt ». À l'impératif présent, le verbe « aller » s'écrit « va » sans « s » (sauf devant « y » ou « en » avec trait d'union).
- `casual_03#1` Registre (mineur) : « Votez en répondant 1 ou 2 ! On verra où on atterrit selon la majorité. Hâte de vous voir jeudi ! » → « Vote en répondant 1 ou 2 ! / Votez en répondant 1 ou 2 ! Hâte de te voir jeudi ! ». La consigne demandait un tutoiement (« informal (tu) ») ; même si l'adresse au groupe justifie le pluriel (« Votez », « vous voir »), une formulation au tutoiement singulier était explicitement attendue par la consigne de registre.

## Fiabilité et vitesse

| Modèle | Textes OK / attendus | Appels en échec / appels | Coupés à la limite | Raisonnement dans la réponse | Latence p50 | Latence p90 | Tokens de sortie | Tokens de raisonnement |
|---|---|---|---|---|---|---|---|---|
| Qwen 3.8 27B | 96 / 96 | 0 / 96 | 0 | 0 | 1,9 s | 4,3 s | 28 691 | 0 |
| DeepSeek V4.1 Flash | 96 / 96 | 0 / 96 | 0 | 0 | 1,3 s | 2,5 s | 25 966 | 0 |
| GLM 5.3 Flash | 96 / 96 | 0 / 96 | 0 | 0 | 3,0 s | 8,9 s | 35 941 | 152 |
| Gemma 4 31B | 96 / 96 | 0 / 96 | 0 | 0 | 7,4 s | 1 min 18 s | 24 770 | 0 |
| GPT-6 Luna | 96 / 96 | 0 / 96 | 0 | 0 | 2,5 s | 4,5 s | 21 761 | 0 |
| Mistral Large 4 | 96 / 96 | 0 / 96 | 0 | 0 | 5,7 s | 11,5 s | 46 694 | 25 022 |

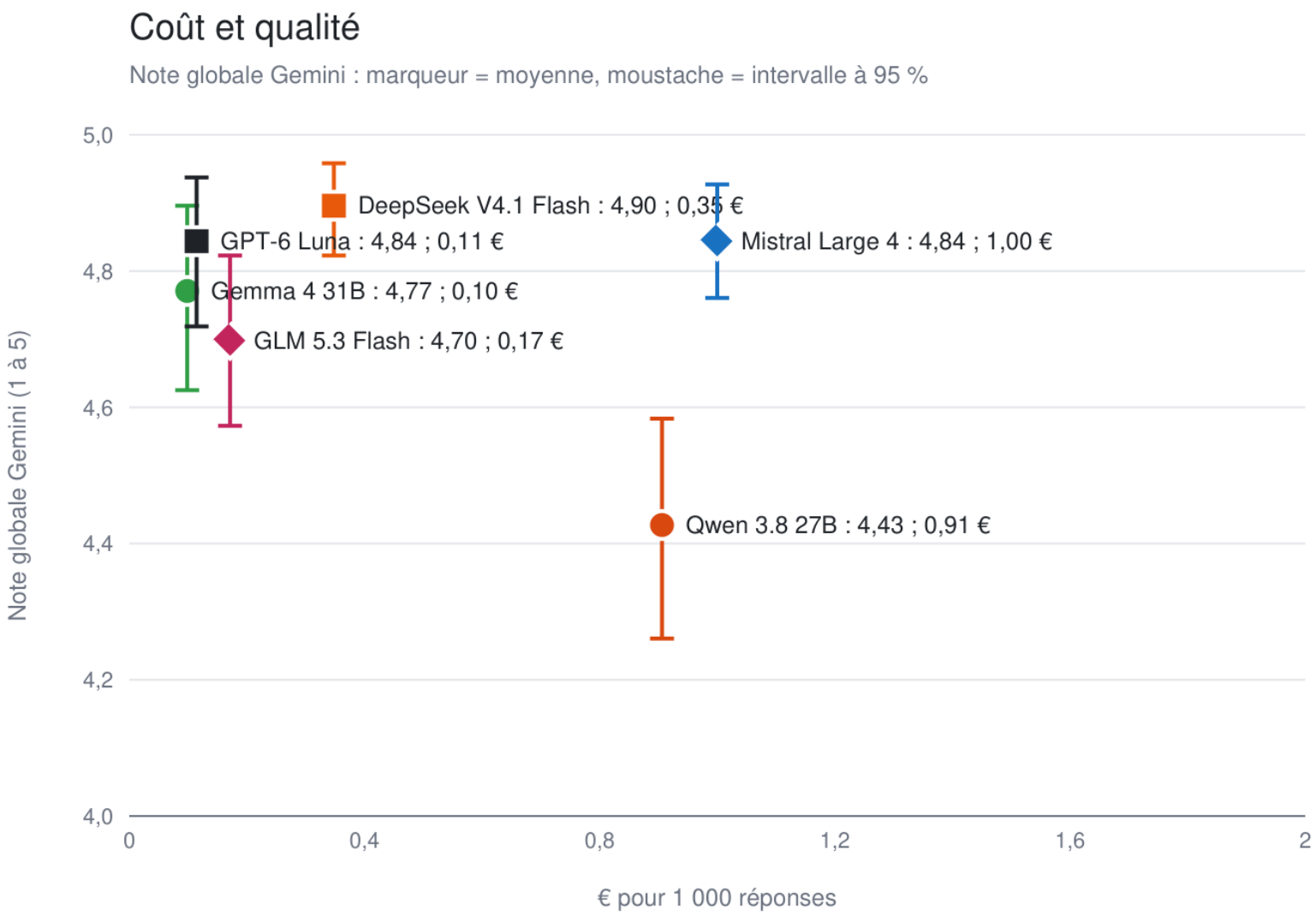
## Coût

Coût du run par compte, en euros (appels en échec compris).

| Compte | EUR |
|---|---|

| | |
|---|---|
| Google (Gemini) | 0,663 |
| OVHcloud AI Endpoints | 0,120 |
| Mistral AI | 0,096 |
| TypeSafe (Jev) | 0,042 |
| Melious | 0,026 |
| OpenAI | 0,011 |
| **Total** | **0,958** |

| **Modèle** | **EUR pour 1 000 réponses** |
|---|---|
| Qwen 3.8 27B | 0,91 |
| DeepSeek V4.1 Flash | 0,35 |
| GLM 5.3 Flash | 0,17 |
| Gemma 4 31B | 0,10 |
| GPT-6 Luna | 0,11 |
| Mistral Large 4 | 1,00 |




## Méthode

Objectif : mesurer la qualité du français que chaque modèle écrit directement (sans traduction), sur des tâches proches de l'usage professionnel d'un assistant, et vérifier si les modèles open-weight chinois écrivent un français moins bon que les autres.

**Corpus**

- 48 consignes rédigées en français pour le benchmark, en 8 styles ; les entreprises, personnes, produits et chiffres sont fictifs, pour qu'aucun modèle ne réponde de mémoire.
- Chaque consigne fixe un registre attendu et une longueur cible. Les extraits du premier style sont volontairement en anglais : c'est en répondant en français à partir de sources anglaises qu'apparaissent anglicismes et calques.
- Le texte intégral de chaque consigne figure en annexe D.

| Style | Consignes | Registres | Longueur cible (mots) |
|---|---|---|---|
| Réponse à partir d'extraits en anglais | 6 | professionnel (vouvoiement) | 170 - 200 |
| Réponse à partir d'extraits en français | 6 | professionnel (vouvoiement) ; neutre, grand public | 160 - 180 |
| E-mail ou courrier professionnel | 6 | soutenu (vouvoiement, formules de courrier) ; professionnel (vouvoiement) | 150 - 220 |
| Synthèse d'un document long | 6 | professionnel (vouvoiement) ; neutre, grand public | 180 - 220 |
| Contenu éditorial | 6 | professionnel (vouvoiement) ; neutre, grand public | 180 - 420 |
| Explication pédagogique | 6 | professionnel (vouvoiement) ; neutre, grand public ; français simple | 250 - 300 |
| Échange informel | 6 | familier (tutoiement) | 50 - 80 |
| Reformulation | 6 | soutenu (vouvoiement, formules de courrier) ; professionnel (vouvoiement) ; neutre, grand public ; français simple | 70 - 120 |

### Génération

- Chaque modèle répond 2 fois à chaque consigne, avec le même prompt système court (annexe A) et la consigne comme unique message utilisateur.
- Température par défaut du fournisseur (aucune n'est envoyée, comme en production), budget de sortie de 8 192 tokens, chaque modèle au réglage de raisonnement le plus bas que son fournisseur respecte (tableau Modèles) ; un raisonnement laissé dans la réponse (blocs `<think>`) est retiré avant l'évaluation.
- Une réponse vide ou coupée par la limite de tokens compte comme un échec. Les limites de débit, erreurs serveur et réseau sont relancées, jusqu'à 5 tentatives.

### Contrôles automatiques

Appliqués à chaque texte, hors code, URL et adresses e-mail.

- Autres alphabets : chaque caractère est classé d'après son nom Unicode. Les caractères chinois, japonais et coréens (ponctuation pleine chasse comprise), cyrilliques, arabes et des autres écritures non latines comptent comme des fuites. Les lettres grecques sont comptées à part et ne sont pas des fuites (π, Δ, µ sont légitimes en français technique), pas plus que les signes à base latine (le « $^{e}$ » de « 1$^{er}$ », les exposants).
- Anglais : lingua découpe le texte en passages par langue ; les passages en anglais d'au moins 4 mots sont comptés. Un texte « contient de l'anglais » quand ils dépassent 5 % de ses caractères.
- LanguageTool (API publique, français) : fautes d'orthographe et autres alertes, pour 1 000 mots. Une faute signalée sur un mot présent dans la consigne est ignorée (noms fictifs). L'API publique classe ses règles de grammaire françaises comme non catégorisées : elles sont comptées avec les autres alertes.
- Grammalecte 2.3 (local, déterministe, français) : alertes de grammaire (accords, conjugaison, participes passés, modes, confusions, élisions, traits d'union) et mots inconnus, pour 1 000 mots, paragraphe par paragraphe. Les mots de la consigne, les sigles et les mots contenant un chiffre ne sont pas des fautes ; les

mots anglais, si ; une alerte de grammaire sur un mot de la consigne écrit avec une majuscule n'est pas comptée (il ignore le genre d'un nom d'entreprise fictif). Ses alertes de typographie (apostrophe typographique, espaces insécables, tirets de liste Markdown) portent sur des conventions qu'on n'attend pas d'un texte Markdown généré : elles restent dans les données mais pas dans les tableaux.

- Typographie française : absence d'espace avant « ; : ! ? » (sauf entre deux chiffres, comme 10:30), guillemets droits au lieu de « », nombres au format anglais non ambigu (1,234,567, 1,234.5, 3.5 %).
- Longueur : nombre de mots de la réponse rapporté à la cible.

**Juges**

Les deux juges notent en aveugle : ils ne voient jamais le nom du modèle ni du fournisseur, et reçoivent la même entrée pour chaque modèle.

- TypeSafe Jev, référence neutre : trois questions fermées sur une échelle de 1 à 5, dont la note est l'espérance de la réponse (moyenne de l'échelle pondérée par les probabilités de Jev). La fluidité reprend mot pour mot la question du benchmark de traduction, pour que les deux benchmarks se comparent ; le naturel traque calques et anglicismes ; le registre voit la consigne et le registre demandé. Questions et échelles en annexe B.
- `gemini-3.8-flash` au niveau de réflexion `low`, avec un schéma de réponse JSON : il reçoit la consigne, le registre et la longueur attendus, puis la réponse ; il donne cinq notes entières de 1 à 5 (correction, naturel, registre, typographie, note globale) et liste chaque problème relevé avec sa catégorie, sa gravité, l'extrait exact et sa correction. Un extrait introuvable dans le texte reste compté mais n'est jamais cité en exemple. Instructions en annexe B.
- Neutralité : Gemini est un modèle Google, comme Gemma 4. Pour chaque texte, on calcule l'écart entre la note globale de Gemini et la fluidité de Jev ; Gemini est soupçonné de favoriser Gemma quand l'écart de Gemma dépasse l'écart moyen des autres modèles sur les mêmes consignes de plus de 0,1 point, avec un intervalle à 95 % qui exclut zéro.

**Classement humain**

Prévu par la méthode mais non réalisé pour ce run : les résultats reposent sur les contrôles automatiques et les deux juges.

**Statistiques et coûts**

- Chaque consigne pèse autant : les générations d'une consigne sont d'abord moyennées, puis l'intervalle à 95 % est un bootstrap par percentiles (2 000 tirages) sur les consignes.
- Deux modèles se comparent par leur différence appariée sur les consignes traitées par les deux ; elle est significative quand son intervalle exclut zéro. La variabilité est la différence absolue moyenne entre deux générations d'une même consigne.
- Les coûts viennent des tokens facturés par chaque fournisseur et des prix de l'annexe C ; les USD sont convertis à 0,86 EUR.

**Limites**

- Gemini est un modèle Google, comme Gemma ; le rapport le compare à Jev, référence neutre.
- LanguageTool signale des noms propres rares et certains choix de style ; seules les fautes sur des mots absents de la consigne sont comptées.

- Grammalecte prend certaines appositions de noms pour des fautes d'accord (« fiche client », « satisfaction utilisateur ») : ses décomptes de grammaire comprennent quelques fausses alertes, réparties sur tous les modèles ; la feuille Grammalecte du classeur les liste.
- Les fournisseurs peuvent servir des poids quantifiés ; les notes décrivent chaque modèle tel que servi par le fournisseur indiqué.

### Annexe A : prompt système

```
Tu es un assistant professionnel. Réponds toujours en français, dans le registre et la longueur demandés, sans préambule ni commentaire sur ta réponse.
```

### Annexe B : instructions des juges

```
You are a demanding French copy editor and linguist. You review a French text written by an AI assistant in answer to a request. You judge the quality of its French
only, not whether its facts are correct.

Score each criterion from 1 (very poor) to 5 (excellent), as integers:
- correctness: grammar, spelling, agreement, conjugation, punctuation.
- naturalness: reads as if written directly in French by a native professional; no calques, no unnecessary anglicisms, no invented or misused words, no words in
another language or script.
- register_fit: tone and level of language fit the request (tu or vous as expected and consistent, formulas, audience).
- typography: French typographic conventions (no-break space before ; : ! ?, « » quotes, number formats, capitals).
- overall: overall quality of the French for a demanding French reader.

List every problem you find in "feedback", most serious first, at most 15 items:
- category: grammar, spelling, anglicism, calque, register, typography, foreign_script (characters of another script, such as Chinese or Cyrillic) or other;
- severity: major (an error a careful reader would notice and that harms credibility) or minor;
- excerpt: the exact words of the text, copied verbatim (at most 12 words);
- correction: the corrected wording, in French;
- explanation: one short sentence, in French.
An empty list is correct when the text has no problem. Do not report a matter of taste as an error.

Then write "comment": one sentence in French giving your overall assessment.
```

Jev `fluency` (french) : The state holds a French business text in Markdown. Rate the quality of its French only: grammar, spelling, natural and idiomatic phrasing, consistent register, and no English words or sentences that a French writer would have translated. Ignore whether the content is accurate, and ignore the Markdown markup.

- 1 : Not French, or unreadable.
- 2 : Hard to read: frequent grammar errors or literal, unnatural phrasing.
- 3 : Understandable but clearly translated: awkward phrasing, anglicisms or English words left in.
- 4 : Natural and correct, with minor awkwardness.
- 5 : Reads as if written directly in French by a professional writer.

Jev `naturalness` (french) : The state holds a French text written by an assistant. Rate how natural it sounds to a native French reader: does it read as if it had been

written directly in French, or as a translation? Penalize calques from English (literal structures, false friends, unidiomatic collocations), anglicisms where a common French word exists, invented or misused words, and words in another language or script. Ignore the content's accuracy and the Markdown markup.

- 1 : Not French, or so unnatural that it is hard to understand.
- 2 : Reads like a poor translation: many calques, anglicisms or invented words.
- 3 : Understandable but clearly not native: several calques or awkward collocations.
- 4 : Natural, with one or two slightly foreign-sounding phrases.
- 5 : Fully idiomatic: indistinguishable from a text written by a native French professional.

Jev `register` (request) : The state holds a request written in French, the register it asks for, and the French answer of an assistant. Rate only whether the answer's tone and register fit the request: formal or informal address (vous or tu) used consistently as expected, level of language, politeness formulas, warmth or firmness asked for, and suitability for the stated audience. Ignore the accuracy of the content and the length.

- 1 : Wrong register throughout (e.g. familiar where a formal letter was asked, or the reverse).
- 2 : Register mostly inappropriate, or address switching between tu and vous.
- 3 : Acceptable register with noticeable lapses in tone or formulas.
- 4 : Appropriate register with minor lapses.
- 5 : Register and tone exactly as requested, consistent from start to end.

### Annexe C : prix

| Service | Entrée / sortie par M tokens | Source | Relevé le |
|---|---|---|---|
| Qwen 3.8 27B | 0,470 / 3,190 USD | Tarifs OVHcloud Public Cloud, AI Endpoints : Qwen3.8-27B | 2026-11-24 |
| DeepSeek V4.1 Flash | 0,310 / 1,250 USD | Tarifs OVHcloud Public Cloud, AI Endpoints : DeepSeek-V4.1-Flash-1017 | 2026-11-24 |
| GLM 5.3 Flash | 0,100 / 0,400 EUR | Page de tarifs Melious : GLM 5.3 Flash (HT) | 2026-11-24 |
| Gemma 4 31B | 0,100 / 0,300 EUR | Page de tarifs Melious : Gemma 4 31B (HT) | 2026-11-24 |
| GPT-6 Luna | 0,100 / 0,500 USD | Tarifs de l'API OpenAI : gpt-6-luna, offre standard | 2026-11-24 |
| Mistral Large 4 | 0,700 / 2,100 USD | Tarifs Mistral AI : Mistral Large 4 | 2026-11-24 |
| Gemini 3.8 Flash | 0,750 / 3,750 USD | Tarifs de l'API Gemini : gemini-3.8-flash, prix de lancement (jusqu'au 31/12/2026, puis 1,50 / 7,50) | 2026-11-24 |
| Jev | 0,042 / 0,000 USD | Tarifs TypeSafe Jev : tokens d'entrée seulement, sortie gratuite | 2026-11-24 |

### Annexe D : consignes

#### Réponse à partir d'extraits en anglais

`rag_en_01` : registre professionnel (vouvoiement), cible 180 mots.

```
Réponds à la question d'un salarié en t'appuyant uniquement sur les extraits ci-dessous.

Question : « Combien de jours de télétravail ai-je le droit de prendre, et que se passe-t-il si je travaille depuis l'étranger ? »

Extraits (politique interne de Tellurion Systems) :
> Section 4.2 - Remote work allowance. Full-time employees who have completed their probation period may work remotely up to three days per week, subject to their manager's approval. Team leads may require on-site presence for onboarding sessions, quarterly planning and client workshops.
> Section 4.5 - Working from abroad. Remote work from another country is limited to 20 working days per calendar year and must be requested at least two weeks in advance through the HR portal. Employees remain responsible for complying with local immigration and tax rules. Requests for countries outside the European Economic Area require written approval from the Legal team.
> Section 4.7 - Equipment. The company provides a laptop and a monitor. A monthly allowance of EUR 30 covers internet costs for employees who work remotely at least two days per week.
```

`rag_en_02` : registre professionnel (vouvoiement), cible 200 mots.

```
Un client demande comment renvoyer un appareil défectueux. Rédige la réponse du service client à partir des extraits suivants.

Question du client : « Mon purificateur d'air Optelia AP-40 fait un bruit de claquement depuis une semaine. Il a 14 mois. Comment le faire réparer ? »

Extraits (conditions de garantie Optelia) :
> Warranty coverage. All Optelia air purifiers are covered by a 24-month limited warranty from the date of purchase. The warranty covers manufacturing defects in materials and workmanship. It does not cover filters, damage caused by misuse, or units that have been opened by unauthorized repairers.
> Return Merchandise Authorization (RMA). Before shipping any unit, customers must open an RMA ticket on support.optelia.example with their serial number and proof of purchase. Our team reviews tickets within two business days and emails a prepaid shipping label.
> Turnaround. Repaired or replaced units are shipped back within 10 business days after reception at our repair center. If a unit cannot be repaired, it is replaced with a new or refurbished unit of equal or better specifications.
> Troubleshooting first. Many noise issues are caused by a misaligned filter. Customers should remove and reinstall the filter, then run the unit on speed 1 for ten minutes before opening a ticket.
```

`rag_en_03` : registre professionnel (vouvoiement), cible 200 mots.

```
Un membre du comité de direction te demande une synthèse en français des résultats trimestriels, à partir de l'extrait du rapport ci-dessous. Réponds à sa question.

Question : « En quelques phrases, comment s'est passé le trimestre et quels sont les points de vigilance ? »

Extrait (Calvane Mobility, Q3 earnings release, fictional) :
> Revenue for the third quarter reached EUR 84.2 million, up 11% year over year, driven by strong demand for fleet electrification services in Germany and the Benelux. Recurring software revenue grew 23% and now accounts for 31% of total revenue.
> Gross margin declined by 1.8 points to 38.4%, mainly due to higher battery procurement costs and one-off logistics expenses related to the relocation of our Lyon warehouse.
> Operating cash flow was negative at EUR -3.1 million, reflecting a temporary build-up of inventory ahead of the winter season. We expect working capital to normalize by the end of Q4.
> Outlook: we confirm our full-year revenue guidance of EUR 330 to 340 million but now expect the EBITDA margin at the lower end of the 9-11% range.
```

`rag_en_04` : registre professionnel (vouvoiement), cible 170 mots.

```
Explique à un collaborateur ce qu'il doit faire, en t'appuyant sur les extraits de la politique de sécurité informatique.
```

```
Question : « On m'a dit que je devais activer un truc de double authentification et changer mon mot de passe. C'est obligatoire ? Et je fais comment avec mon téléphone perso ? »

Extraits (Fontbrune Group – Information Security Policy) :
> 2.1 Passwords must be at least 14 characters long. Passphrases are encouraged. Passwords must not be reused across company and personal accounts. Mandatory rotation has been discontinued; passwords must be changed only when a compromise is suspected.
> 2.3 Multi-factor authentication (MFA) is mandatory for all employees by March 31. The approved method is the Fontbrune Authenticator app. SMS codes are no longer accepted.
> 2.4 Personal devices may be used for MFA only. The authenticator app does not give the company access to personal data on the device. Employees who do not wish to use a personal phone can request a hardware security key from the IT desk.
```

`rag_en_05` : registre professionnel (vouvoiement), cible 190 mots.

```
Un technicien de maintenance te pose une question. Réponds à partir des extraits du manuel constructeur.

Question : « À quelle fréquence faut-il remplacer les garnitures mécaniques de la pompe HX-220 et qu'est-ce qui doit m'alerter entre deux révisions ? »

Extraits (Kestrel Pumps – HX-220 Maintenance Manual) :
> 7.1 Scheduled maintenance. Inspect the mechanical seal every 2,000 operating hours. Replace the seal every 8,000 operating hours or every two years, whichever comes first. In abrasive fluids, halve these intervals.
> 7.2 Warning signs. Visible leakage of more than 10 drops per minute at the seal, unusual vibration above 4.5 mm/s RMS, or a bearing temperature exceeding 80 °C indicate premature wear. Stop the pump and schedule an inspection.
> 7.3 Dry running. Never run the pump dry, even for a few seconds. Dry running destroys the seal faces and voids the warranty.
> 7.4 Spare parts. Use only genuine seal kits (ref. HX-SK-220). Third-party kits may not withstand the rated pressure of 16 bar.
```

`rag_en_06` : registre professionnel (vouvoiement), cible 200 mots.

```
Un client B2B conteste une facture après une panne. Rédige une réponse claire en t'appuyant sur les extraits du contrat.

Question du client : « Votre plateforme a été indisponible environ 5 heures le mois dernier. Avons-nous droit à un dédommagement ? »

Extraits (Lumeris Cloud – Service Level Agreement) :
> Availability commitment. Lumeris commits to a monthly availability of 99.9% for the Production tier, measured by our external monitoring service. Scheduled maintenance announced at least 72 hours in advance is excluded.
> Service credits. If monthly availability falls below 99.9%, the customer is entitled to a credit of 10% of the monthly fee; below 99.5%, 25%; below 99.0%, 50%. Credits are capped at 50% of the monthly fee.
> Claim procedure. Credit requests must be submitted within 30 days after the end of the affected month, through the customer portal, with the dates and times of the incident. Credits are applied to the next invoice and cannot be exchanged for cash.
```

### Réponse à partir d'extraits en français

`rag_fr_01` : registre professionnel (vouvoiement), cible 170 mots.

```
Réponds à la question d'une salariée en t'appuyant uniquement sur les extraits de l'accord d'entreprise.
```

```
Question : « Je voudrais poser trois semaines en août et une semaine à Noël. Est-ce possible, et avant quand dois-je faire ma demande ? »

Extraits (accord sur les congés de la société Brasserie des Quatre Vents) :
> Article 3 – Période de prise du congé principal. Le congé principal de quatre semaines est pris entre le 1er mai et le 31 octobre. Au moins deux semaines consécutives doivent être posées pendant cette période.
> Article 5 – Délais de prévenance. Les demandes portant sur la période estivale sont déposées avant le 1er mars. Les autres demandes sont déposées au moins un mois avant la date de départ. Le responsable répond sous quinze jours ; à défaut, la demande est réputée acceptée.
> Article 6 – Fermeture de fin d'année. Le site de production est fermé entre Noël et le jour de l'An. Ces jours sont décomptés de la cinquième semaine de congés payés.
```

### rag_fr_02 : registre professionnel (vouvoiement), cible 160 mots.

```
Un commercial demande s'il peut se faire rembourser certaines dépenses. Réponds à partir des extraits de la politique de notes de frais.

Question : « J'ai invité un client à dîner à 95 euros pour deux, j'ai pris un taxi depuis la gare et j'ai payé un parking à l'aéroport pendant trois jours. Tout est remboursé ? »

Extraits (politique de notes de frais de Verdane Énergie) :
> Repas d'affaires : plafond de 45 € TTC par convive, sur présentation de la facture détaillée et du nom des invités. Tout dépassement reste à la charge du collaborateur, sauf accord écrit préalable du directeur commercial.
> Déplacements locaux : les taxis et VTC sont remboursés lorsque aucun transport en commun n'est disponible dans un délai raisonnable ou après 21 h.
> Stationnement : les parkings d'aéroport sont remboursés dans la limite de 20 € par jour. Le parking longue durée doit être privilégié.
```

### rag_fr_03 : registre neutre, grand public, cible 180 mots.

```
Un particulier pose une question sur son contrat d'assurance habitation. Réponds-lui à partir des extraits des conditions générales.

Question : « Un dégât des eaux venant de chez mon voisin du dessus a abîmé mon parquet. Je suis locataire. Qui paie et que dois-je faire ? »

Extraits (conditions générales Habitation Sérénité de Mutuelle Horizon) :
> Déclaration du sinistre : l'assuré déclare le sinistre dans les cinq jours ouvrés suivant sa découverte, par l'espace adhérent ou par téléphone. Il joint un constat amiable de dégât des eaux rempli avec le voisin concerné.
> Prise en charge : pour les dégâts inférieurs à 5 000 € HT, l'assureur de l'occupant du logement sinistré indemnise les dommages aux embellissements et aux biens mobiliers, quelle que soit l'origine de la fuite.
> Franchise : une franchise de 150 € reste à la charge de l'assuré.
```

### rag_fr_04 : registre professionnel (vouvoiement), cible 180 mots.

```
Une chargée de compte demande quelles sont les étapes pour démarrer avec un nouveau client. Réponds à partir de la procédure interne.

Question : « J'ai signé un nouveau client hier. Qu'est-ce que je dois faire cette semaine ? »

Extraits (procédure d'intégration client de l'agence Atelier Brume) :
> Étape 1 (J+1) : créer la fiche client dans le CRM, avec le bon de commande signé et les coordonnées du référent.
> Étape 2 (J+2) : envoyer le questionnaire de cadrage et proposer trois créneaux pour la réunion de lancement.
> Étape 3 (J+5 au plus tard) : tenir la réunion de lancement avec le chef de projet et le directeur artistique. Le compte rendu est envoyé au client sous 48 heures.
> Étape 4 : ouvrir l'espace de partage de fichiers et y déposer le calendrier de production validé.
```

> Point d'attention : aucune production ne démarre avant la réception de l'acompte de 30 %.

### rag_fr_05 : registre professionnel (vouvoiement), cible 170 mots.

Un manager absent lors du dernier comité de direction te demande ce qui a été décidé. Réponds à partir du compte rendu.

Question : « Qu'est-ce qui a été décidé sur le déménagement et sur les recrutements ? »

Extrait du compte rendu du comité de direction de Tellurion Systems (12 septembre) :
> Déménagement : le comité valide le transfert du siège à Villeurbanne au deuxième trimestre de l'année prochaine. Le budget de 1,2 M€ est approuvé. Un groupe de travail de six salariés volontaires sera constitué pour l'aménagement des espaces.
> Recrutements : gel des embauches sur les fonctions support jusqu'à la fin de l'année. Les quatre postes d'ingénieurs prévus sont maintenus. Le poste de directeur financier sera pourvu en interne par intérim.
> Points reportés : la révision de la politique de télétravail est reportée au comité d'octobre.

### rag_fr_06 : registre neutre, grand public, cible 160 mots.

Un client d'une boutique en ligne veut retourner un article. Réponds à partir des conditions générales de vente.

Question : « J'ai reçu un manteau il y a dix jours, il ne me va pas. Je peux le renvoyer ? C'est gratuit ? »

Extraits (conditions générales de vente de la boutique Lin & Laine) :
> Droit de rétractation : le client dispose de 14 jours à compter de la réception de sa commande pour exercer son droit de rétractation, sans avoir à justifier de motif.
> Conditions de retour : l'article est retourné non porté, avec ses étiquettes, dans son emballage d'origine.
> Frais de retour : les retours sont gratuits en point relais. Un retour par transporteur à domicile est facturé 6,90 €.
> Remboursement : il intervient dans un délai de 14 jours après réception du retour, par le moyen de paiement utilisé lors de la commande.

## E-mail ou courrier professionnel

### email_01 : registre professionnel (vouvoiement), cible 180 mots.

Rédige la réponse de Claire Morand, responsable du service client de la société Optelia, à un client mécontent. Sa commande de 20 purificateurs d'air pour ses bureaux devait être livrée le 3 et ne l'a pas été ; il menace d'annuler. La livraison est désormais prévue le 10 à cause d'une rupture chez un fournisseur. Claire propose la livraison gratuite et un filtre de rechange offert par appareil. Ton : empathique, professionnel, sans formule creuse.

### email_02 : registre professionnel (vouvoiement), cible 150 mots.

Rédige un e-mail de relance pour une facture impayée. La facture n° 2024-0457 de 3 480 € TTC, émise par le cabinet de conseil Calvane pour une mission de formation, était due il y a 30 jours ; une première relance est restée sans réponse. Le ton est ferme mais courtois. Rappelle les pénalités de retard prévues au contrat et propose un échéancier si le client rencontre des difficultés.

### email_03 : registre professionnel (vouvoiement), cible 200 mots.

Rédige une invitation par e-mail à un séminaire partenaires. Organisateur : Kestrel Pumps. Date : jeudi 12 juin, de 9 h 30 à 17 h, au Domaine de la Source, près d'Annecy. Programme : présentation de la nouvelle gamme de pompes basse consommation, retours d'expérience de deux clients, ateliers techniques l'après-midi, dîner

facultatif. Inscription avant le 20 mai via un formulaire en ligne. Le ton doit donner envie de venir sans être publicitaire.

`email_04` : registre professionnel (vouvoiement), cible 220 mots.

Rédige un e-mail interne de la direction des systèmes d'information de Fontbrune Group annonçant à tous les salariés le remplacement de l'outil de gestion des congés par un nouveau logiciel à partir du 1er avril. Points à couvrir : pourquoi ce changement (outil actuel ancien, plus de saisie double avec la paie), ce qui change pour eux, les sessions de formation de 30 minutes proposées la dernière semaine de mars, le fait que les soldes de congés sont repris automatiquement, et à qui s'adresser en cas de question.

`email_05` : registre soutenu (vouvoiement, formules de courrier), cible 200 mots.

Rédige un courrier formel de résiliation de contrat. Expéditeur : Martine Duclos, directrice administrative de la société Lin & Laine. Destinataire : la société NetPropre, prestataire de nettoyage des locaux. Le contrat n° NP-2021-118 est résilié à son échéance du 30 juin, en respectant le préavis de trois mois prévu à l'article 9. Demande la restitution des badges d'accès et un état des lieux contradictoire. Utilise les conventions d'un courrier recommandé (objet, formule d'appel, formule de politesse).

`email_06` : registre professionnel (vouvoiement), cible 150 mots.

Rédige la réponse à une candidate non retenue pour un poste de cheffe de projet marketing chez Atelier Brume, après deux entretiens. Le choix s'est porté sur un profil ayant davantage d'expérience dans le secteur de la cosmétique. La réponse doit être humaine, valoriser ses qualités (sa présentation du cas pratique a été très appréciée) et proposer de garder sa candidature pour de futurs postes, avec son accord.

### Synthèse d'un document long

`summary_01` : registre professionnel (vouvoiement), cible 200 mots.

Fais la synthèse de la transcription de réunion ci-dessous pour les personnes absentes : décisions, actions avec responsables et échéances, points ouverts. Utilise des listes.

Transcription (réunion projet « Portail fournisseurs », Verdane Énergie) :
Sophie : Bon, on commence. Le point principal, c'est le retard sur l'intégration avec l'ERP. Karim, tu peux faire le point ?
Karim : Oui. On a terminé les flux de commandes, mais les flux de factures bloquent toujours. Le prestataire ERP nous a envoyé une nouvelle version de l'API seulement lundi, et elle ne gère pas les avoirs. Sans les avoirs, on ne peut pas ouvrir le portail aux fournisseurs de la catégorie maintenance, qui en émettent beaucoup.
Sophie : Donc qu'est-ce qu'on propose ?
Karim : Deux options. Soit on attend la version corrigée, annoncée pour fin du mois, soit on ouvre d'abord le portail aux fournisseurs de fournitures de bureau, qui n'ont quasiment pas d'avoirs.
Élodie : Côté achats, je préfère l'ouverture progressive. Les fournisseurs de fournitures sont demandeurs, et ça nous permettra de tester le support.
Sophie : D'accord, on part sur l'ouverture progressive. Élodie, tu peux préparer la liste des fournisseurs de la première vague ?
Élodie : Oui, pour vendredi prochain. Il y en a une quarantaine.
Sophie : Autre sujet, la formation. Marc, où en est-on ?
Marc : Le support de formation est prêt. Je propose deux webinaires d'une heure, un mardi matin et un jeudi après-midi, pour que les fournisseurs puissent choisir.
Sophie : Parfait. Et les comptes utilisateurs ?
Karim : On peut les créer en masse à partir d'un fichier, mais il faut que les adresses e-mail soient à jour. Aujourd'hui, environ 15 % sont manquantes ou invalides dans la base.
Élodie : Je peux demander à mon équipe de les vérifier en même temps que la liste.
Sophie : Très bien. Dernier point : le budget. On a consommé 70 % du budget pour 60 % du périmètre. Je vais devoir en parler au comité de pilotage du 18. Karim,

```
j'aurai besoin d'une estimation du reste à faire avant le 15.
Karim : Ça marche.
Sophie : On n'a pas tranché la question de l'assistance en anglais pour les fournisseurs étrangers. On la garde pour la prochaine fois. Merci à tous.
```

### `summary_02` : registre professionnel (vouvoiement), cible 220 mots.

```
Résume ce rapport d'incident pour la direction générale, qui n'est pas technique : ce qui s'est passé, l'impact pour les clients, la cause, les mesures prises et prévues. Pas de jargon.

Rapport post-incident (Lumeris Cloud, incident du 4 février) :
Chronologie. À 09 h 12, l'équipe d'astreinte reçoit des alertes de latence sur le service de stockage de la région Paris. À 09 h 25, les premiers clients signalent des erreurs lors de l'envoi de fichiers. À 09 h 40, l'incident est déclaré majeur. À 10 h 05, l'équipe identifie un remplissage anormal des disques de journalisation sur trois serveurs de métadonnées. À 10 h 30, les journaux sont purgés et le service se rétablit progressivement. À 11 h 02, le service est revenu à la normale.
Impact. Pendant 1 h 50, environ 18 % des requêtes d'envoi de fichiers ont échoué pour les clients hébergés dans la région Paris. Aucune donnée n'a été perdue. Les téléchargements n'ont pas été affectés. 212 tickets de support ont été ouverts.
Cause racine. Une mise à jour déployée le 2 février a augmenté le niveau de détail des journaux sur les serveurs de métadonnées. La rotation automatique des journaux n'était pas configurée sur ces serveurs, car ils avaient été installés avec un modèle différent des autres. Les disques se sont remplis en 48 heures.
Mesures immédiates. Purge des journaux, retour au niveau de détail précédent, vérification de la rotation sur l'ensemble du parc.
Mesures prévues. Alerte sur le taux de remplissage des disques à 70 % (au lieu de 90 %) ; harmonisation des modèles d'installation d'ici fin mars ; revue obligatoire des changements de configuration des journaux ; communication proactive aux clients dès la déclaration d'un incident majeur, alors que la première communication n'est partie qu'à 10 h 15.
```

### `summary_03` : registre professionnel (vouvoiement), cible 200 mots.

```
Synthétise les enseignements de ces entretiens utilisateurs en cinq points maximum, chacun appuyé par ce que disent les personnes interrogées.

Notes d'entretiens (application de covoiturage domicile-travail de Calvane Mobility, 6 entretiens) :
- Nadia, 34 ans, infirmière : horaires décalés, ne trouve jamais de trajet compatible. « L'appli me propose des trajets à 8 h, moi je commence à 6 h 30. » A désinstallé après deux semaines.
- Thomas, 41 ans, comptable : utilisateur régulier, trois fois par semaine. Apprécie de partager les frais. Trouve la messagerie peu pratique : « On s'envoie des SMS en parallèle parce que les notifications arrivent en retard. »
- Inès, 27 ans, développeuse : a peur de tomber sur des inconnus. « J'aimerais voir si la personne travaille vraiment dans ma boîte. » Utilise l'appli seulement avec des collègues qu'elle connaît.
- Gérard, 58 ans, technicien : trouve l'inscription trop longue, abandonnée à l'étape de la photo du permis.
- Lucie, 31 ans, juriste : utilisatrice satisfaite, mais regrette de ne pas pouvoir réserver un trajet régulier pour toute la semaine en une fois.
- Samir, 45 ans, chef d'équipe : conducteur, aimerait être sûr d'être payé. « Deux fois, des passagers ne sont pas venus et je n'ai rien eu. »
```

### `summary_04` : registre professionnel (vouvoiement), cible 200 mots.

```
Résume en français la note d'audit ci-dessous pour le comité d'audit : principaux constats, niveau de risque, recommandations.

Internal audit memo (Fontbrune Group, procurement process, fictional):
Scope. The audit covered purchase orders issued between January and June across the three French subsidiaries, a sample of 120 orders out of 4,310.
Findings. 1) In 22 of the 120 orders sampled, the purchase order was created after the supplier invoice had been received, which bypasses the budget approval step. 2) Eleven suppliers were created in the master data without the mandatory bank account verification call. 3) The approval thresholds in the ERP do not match the delegation of authority policy updated last year: orders between EUR 50,000 and EUR 100,000 require only one approval instead of two. 4) Positive point: contract renewals are well tracked, with no expired contract still in use.
```

Risk assessment. Findings 2 and 3 are rated high risk (fraud exposure); finding 1 is rated medium.
Recommendations. Block invoice registration without a prior purchase order, except for a short list of exempted categories; enforce the bank verification call with a mandatory field in the supplier creation workflow; align the ERP thresholds with the delegation policy before the end of the quarter.

`summary_05` : registre professionnel (vouvoiement), cible 200 mots.

Résume cette étude de marché pour un investisseur pressé : taille du marché, tendances, acteurs, opportunité pour l'entreprise.

Extrait de l'étude (marché français des purificateurs d'air domestiques, cabinet Optima Conseil, données fictives) :
Le marché français des purificateurs d'air domestiques a atteint 410 millions d'euros l'an dernier, en hausse de 9 % après un recul en 2022. La croissance est portée par les ménages urbains avec enfants et par les personnes allergiques, qui représentent 42 % des acheteurs. Le prix moyen de vente a augmenté de 6 %, signe d'une montée en gamme vers des appareils connectés.
Trois acteurs dominent avec 55 % des ventes en valeur, tous d'origine étrangère. Les marques françaises représentent moins de 10 % du marché, mais bénéficient d'une image de qualité et de réparabilité appréciée : 61 % des acheteurs déclarent être prêts à payer plus cher un appareil réparable.
Le principal frein reste le coût des filtres de rechange, jugé trop élevé par deux acheteurs sur trois. La vente de filtres par abonnement progresse rapidement, mais reste peu proposée.
Les canaux en ligne représentent 58 % des ventes. La grande distribution spécialisée recule.
Opportunité : un acteur français positionné sur la réparabilité, avec une offre d'abonnement aux filtres, pourrait viser 5 % du marché en trois ans.

`summary_06` : registre neutre, grand public, cible 180 mots.

Résume cette note juridique pour des managers non juristes : ce qui change, à partir de quand, ce qu'ils doivent faire concrètement.

Note du service juridique (Brasserie des Quatre Vents, sur la nouvelle procédure de signalement interne) :
À compter du 1er septembre, l'entreprise met en place un dispositif de recueil des signalements conforme aux obligations applicables aux entreprises de plus de cinquante salariés. Tout salarié, ancien salarié, candidat, prestataire ou fournisseur peut signaler des faits constitutifs d'un crime, d'un délit, d'une violation grave du droit ou d'une menace pour l'intérêt général dont il a eu connaissance dans le cadre de ses activités.
Le signalement est effectué via une plateforme externe garantissant la confidentialité de l'identité de l'auteur, des personnes visées et des informations recueillies. Le lanceur d'alerte peut également s'adresser directement aux autorités compétentes.
L'auteur du signalement est informé de la réception sous sept jours, puis des suites données sous trois mois.
Aucune mesure de représailles ne peut être prise à l'encontre d'un lanceur d'alerte de bonne foi. Les managers qui reçoivent un signalement oral doivent orienter la personne vers la plateforme et s'abstenir de mener eux-mêmes une enquête.

### Contenu éditorial

`editorial_01` : registre neutre, grand public, cible 420 mots.

Rédige un article de blog pour le site d'Optelia intitulé « Cinq idées reçues sur la qualité de l'air intérieur ». Public : particuliers. Le ton est clair, vivant et rigoureux, sans alarmisme. Structure : une introduction, cinq intertitres, une conclusion. N'invente pas de chiffres précis ni de sources.

`editorial_02` : registre professionnel (vouvoiement), cible 180 mots.

Rédige un post LinkedIn pour Inès Garnier, directrice générale de Calvane Mobility, annonçant que l'entreprise vient de franchir le cap des 10 000 salariés utilisant son application de covoiturage domicile-travail. Elle remercie les équipes et les entreprises clientes, et partage un enseignement : le covoiturage marche quand on le rend simple et sûr. Ton : authentique, enthousiaste sans excès, à la première personne. Deux ou trois émojis au plus.

`editorial_03` : registre neutre, grand public, cible 250 mots.

```
Rédige l'éditorial de la newsletter mensuelle de la Brasserie des Quatre Vents, une brasserie artisanale bretonne. Ce mois-ci : la nouvelle bière de saison au sarrasin, la réouverture de la salle de dégustation après travaux et la participation au marché de Noël de Quimper. Ton chaleureux et complice avec les lecteurs, sans tutoiement.
```

`editorial_04` : registre professionnel (vouvoiement), cible 350 mots.

```
Rédige un communiqué de presse annonçant le lancement de la pompe HX-300 de Kestrel Pumps, qui consomme 30 % d'énergie en moins que la génération précédente selon les essais internes de l'entreprise. Disponible à partir de septembre dans toute l'Europe. Inclue un titre, un chapeau, une citation du directeur technique, Paul Le Goff, et un paragraphe « À propos de Kestrel Pumps ».
```

`editorial_05` : registre neutre, grand public, cible 400 mots.

```
Rédige une tribune pour la presse économique, signée par la dirigeante d'une PME industrielle, défendant l'idée que la semaine de quatre jours n'est pas réservée aux entreprises du numérique. Elle s'appuie sur l'expérience de son atelier (sans chiffres inventés trop précis), reconnaît les difficultés et appelle au pragmatisme. Ton argumenté, personnel, élégant.
```

`editorial_06` : registre professionnel (vouvoiement), cible 350 mots.

```
Rédige une étude de cas client pour le site de Lumeris Cloud : la mutuelle fictive Mutuelle Horizon a migré son espace adhérent vers la plateforme Lumeris. Structure : le contexte et les difficultés, la solution, les résultats (temps de chargement divisé par trois, plus aucune interruption lors des pics de fin de mois), une citation du DSI, Yann Prigent. Ton factuel et convaincant.
```

### Explication pédagogique

`explainer_01` : registre français simple, cible 250 mots.

```
Explique ce qu'est l'inflation à un adolescent de 15 ans qui n'a jamais étudié l'économie : pourquoi les prix augmentent, ce que ça change pour lui, et pourquoi les banques centrales s'en occupent. Utilise un exemple concret de la vie quotidienne.
```

`explainer_02` : registre neutre, grand public, cible 300 mots.

```
Explique simplement comment fonctionne une pompe à chaleur air-eau à un propriétaire qui envisage d'en installer une : le principe, pourquoi elle consomme moins qu'un chauffage électrique classique, ses limites par grand froid, et les questions à poser à l'installateur.
```

`explainer_03` : registre professionnel (vouvoiement), cible 280 mots.

```
Explique à une directrice commerciale, qui n'est pas technicienne, comment un assistant d'intelligence artificielle peut répondre à des questions à partir des documents internes de l'entreprise : le principe de la recherche dans les documents puis de la rédaction de la réponse, pourquoi il cite ses sources, et pourquoi il peut parfois se tromper.
```

`explainer_04` : registre français simple, cible 300 mots.

```
Explique les bases du RGPD au gérant d'une petite boulangerie qui tient un fichier de clients fidèles et envoie une newsletter : quelles données il peut garder, ce
```

qu'il doit dire à ses clients, combien de temps garder les données, et que faire si un client demande la suppression de ses informations.

`explainer_05` : registre neutre, grand public, cible 250 mots.

Explique les intérêts composés à une personne qui commence à épargner : ce que c'est, un exemple chiffré simple sur plusieurs années, pourquoi commencer tôt compte, et une mise en garde sur les placements qui promettent des rendements élevés.

`explainer_06` : registre neutre, grand public, cible 280 mots.

Explique à des parents comment fonctionne le système immunitaire et à quoi sert un vaccin, avec une analogie compréhensible par tous. Reste factuel, rassurant sans minimiser, et invite à poser les questions médicales à un professionnel de santé.

### Échange informel

`casual_01` : registre familier (tutoiement), cible 60 mots.

Un ami t'écrit : « Salut ! Je suis à Lyon ce week-end avec ma copine, tu aurais une idée de resto sympa pas trop cher ? » Réponds-lui comme à un ami, de façon naturelle et courte. Tu peux proposer un type de cuisine ou un quartier sans citer de nom d'établissement réel.

`casual_02` : registre familier (tutoiement), cible 50 mots.

Une collègue avec qui tu es en tutoiement t'écrit sur la messagerie interne : « Hello, je suis coincée chez le dentiste, tu pourrais décaler notre point de 14 h à 15 h 30 et prévenir Julien ? Merci mille fois !! » Réponds-lui.

`casual_03` : registre familier (tutoiement), cible 70 mots.

Rédige le message que tu postes sur le groupe de discussion de ton équipe pour organiser un afterwork jeudi prochain pour fêter la fin d'un gros projet. Propose deux options de lieu (un bar avec terrasse ou un bowling) et demande à chacun de voter. Ton léger, avec un peu d'humour.

`casual_04` : registre familier (tutoiement), cible 60 mots.

Ton frère t'envoie : « J'ai raté mon permis pour la deuxième fois… j'en ai marre. » Réponds-lui avec bienveillance et un peu d'humour, sans faire la morale.

`casual_05` : registre familier (tutoiement), cible 80 mots.

Un collègue junior, que tu tutoies, te demande ton avis sur la présentation qu'il a faite ce matin en réunion. Elle était claire et bien préparée, mais trop longue, et il lisait beaucoup ses diapositives. Donne-lui un retour franc et encourageant, en quelques phrases.

`casual_06` : registre familier (tutoiement), cible 50 mots.

Écris un petit mot à glisser dans la boîte aux lettres de tes voisins du dessus, que tu connais un peu et tutoies, pour les prévenir que tu organises un anniversaire samedi soir et qu'il risque d'y avoir du bruit jusqu'à minuit. Invite-les à passer boire un verre.

### Reformulation

### rewrite_01 : registre français simple, cible 110 mots.

```
Reformule cette clause en français clair, compréhensible par un client particulier, sans perdre d'information :

« Le Prestataire ne saurait être tenu pour responsable de tout dommage indirect, immatériel ou consécutif, en ce compris, sans que cette liste soit limitative, la perte de chiffre d'affaires, de bénéfice, de clientèle ou de données, résultant de l'utilisation ou de l'impossibilité d'utiliser le Service, quand bien même le Prestataire aurait été informé de l'éventualité de tels dommages, sa responsabilité étant en tout état de cause limitée au montant des sommes effectivement versées par le Client au titre des douze (12) derniers mois précédant le fait générateur. »
```

### rewrite_02 : registre français simple, cible 120 mots.

```
Réécris ce courrier administratif en langage simple, pour un usager qui maîtrise mal le français écrit. Garde toutes les informations utiles (dates, documents, démarche).

« Madame, Monsieur, nous accusons réception de votre demande d'aide au logement. Afin de procéder à l'instruction de votre dossier, il vous appartient de nous faire parvenir, dans un délai de quinze jours à compter de la date du présent courrier, les pièces justificatives listées ci-après : avis d'imposition de l'année N-1, attestation de loyer complétée par votre bailleur, relevé d'identité bancaire à votre nom. À défaut de réception de l'ensemble de ces pièces dans le délai imparti, votre demande ne pourra être instruite et fera l'objet d'un classement sans suite. »
```

### rewrite_03 : registre neutre, grand public, cible 120 mots.

```
Réécris ce paragraphe technique pour une fiche produit destinée au grand public, en gardant les informations importantes :

« Le module HX-300 intègre un variateur de fréquence à commande vectorielle permettant une régulation dynamique du débit en fonction de la pression différentielle mesurée en aval, ce qui réduit la consommation énergétique en charge partielle et limite les coups de bélier lors des phases de démarrage et d'arrêt. L'indice de protection IP55 autorise une installation en local technique humide. »
```

### rewrite_04 : registre soutenu (vouvoiement, formules de courrier), cible 120 mots.

```
Transforme ce message en un courriel formel adressé au directeur général d'une entreprise cliente, M. Henri Vasseur :

« Salut, on a bien reçu votre commande mais y a un souci avec les délais, le transporteur est en grève donc ça sera plutôt la semaine prochaine. On vous tient au courant, désolé pour le contretemps ! »
```

### rewrite_05 : registre professionnel (vouvoiement), cible 70 mots.

```
Rends ce texte plus concis et plus lisible, en deux ou trois phrases, sans perdre d'information :

« Dans le cadre de la démarche d'amélioration continue qui est la nôtre et dans le but de pouvoir répondre au mieux aux attentes qui sont celles de nos clients, nous avons pris la décision, après mûre réflexion et à l'issue d'un certain nombre d'échanges avec les différentes parties prenantes concernées, de procéder à la mise en place, à compter du premier jour du mois de mars, d'un nouveau service d'assistance téléphonique qui sera accessible du lundi au samedi, de huit heures du matin à vingt heures, et ce sans aucun surcoût pour les clients disposant d'un contrat en cours de validité. »
```

### rewrite_06 : registre professionnel (vouvoiement), cible 90 mots.

```
Réécris ce message en bon français professionnel, en remplaçant les anglicismes inutiles par des équivalents français naturels :
```

« Hello la team, suite au call de ce matin, je vais checker le process d'onboarding et forwarder le deck updaté à Julie ASAP. On se fait un meeting en fin de week pour challenger les KPIs et définir les next steps. N'hésitez pas à me ping si besoin, je suis full dispo cet aprem. »