

IA & Archives – Partie Pratique

1 Notez vos observations après chacune des étapes

- Ouvrez l'application de chat fournie.
- Sélectionnez le **Modèle A** et entrez le prompt suivant :

*Je suis archiviste et je travaille sur le fonds Georges Lemaître.
Peux-tu me résumer brièvement qui il était ?*

- Une fois la réponse obtenue, **ne réinitialisez pas la conversation**. Changez simplement le modèle via le menu déroulant pour sélectionner le **Modèle B**. Posez ensuite cette question simple, sans rappeler le sujet :

Quelles sont les dates extrêmes de ses principaux travaux ?

- Passez-maintenant au **Modèle C**, ne réinitialisez pas encore la conversation. Posez-lui maintenant la question suivante :

Quels sont les travaux qui lui ont valu de passer à la postérité ?

- Réinitialisez la conversation (bouton rouge en haut à droite). Recommencez toutes les étapes, mais **réinitialisez la conversation avant chaque changement de modèle**. Que constatez-vous ?

2 Évaluez la qualité des réponses données par chacun des modèles

Pour cet exercice, nous allons comparer la qualité des réponses fournies par les différents modèles lorsqu'on leur fournit **exactement le même prompt**. A cet effet, nous vous demandons de commencer par récupérer les documents suivants :

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_2a.png

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_2b.png

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_2c.txt

Pour chacun de ces documents, vous donnerez ensuite le prompt suivant à chacun des modèles (en y joignant une copie du document téléchargé).

Agis comme un archiviste expérimenté. Analyse le document ci-dessus. Extrais-en : 1) La date exacte, 2) Les parties prenantes (noms et rôles), 3) L'objet de l'acte. Formate la réponse en Markdown.

Comparez les résultats obtenus. Y a t il eu des hallucinations ? Quel modèle a mieux respecté la consigne ? Quel modèle a utilisé le style le plus « archivistique » ? Établissez votre podium.

3 Biais de complaisance vs « Adversarial Prompting »

Objectif : Réaliser que la formulation de votre prompt (le *framing*) influence directement et radicalement la réponse donnée par l'IA et comprendre pourquoi.

Matériel :

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_3.md

Partie A : Le mode "Complaisant"

Utilisez le prompt suivant, pour demander à l'un des modèles de vous donner une réponse :

On me demande de rentrer un brouillon pour le dossier de financement sur lequel je travaille actuellement. Que penses-tu de celui-ci ? A ton avis, j'ai des chances d'être accepté ?

Partie B : Le mode "Critique" (Adversarial)

Videz la conversation, puis envoyez le prompt suivant au même modèle que précédemment en y joignant à nouveau une copie du document que vous avez téléchargé.

Je suis membre d'un comité de financement très sélectif, les projets que nous finançons sont extrêmement onéreux. Une erreur de casting pourrait facilement mettre tout l'organisme de financement en péril. Comme notre budget est serré et que nous ne finançons que l'excellence, j'ai demandé aux candidats de me soumettre un premier brouillon du dossier de financement sur lequel il travaillent afin de pouvoir faire un premier tri sans que personne ne perde trop de temps. Analyse ce dossier sans aucune complaisance. N'hésite pas à chercher la petite bête. Identifie les failles méthodologiques ou me dire pourquoi je devrais rejeter ce projet. Enfin donne moi une critique générale sur ce dossier puis émets une recommandation : (Accept, Weak Reject, Strong Reject).

Conclusions :

Comparez les deux réponses données par le modèle. Comparez aussi la recommandation faite par le même modèle à propos du même document. Quelles conclusions en tirez-vous ? Comment expliquez-vous ce phénomène ?

4 Créez une page web à partir d'un fichier EAC-XML

Commencez par télécharger les ressources nécessaires à réaliser l'exercice:

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_4a.png

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_4b.eac.xml

Dans cet exercice, il vous est demandé de parvenir à réaliser une page web complète, fonctionnelle, et qui respecte le style graphique d'un autre site à partir d'un document EAC-XML. Comment avez vous procédé ?

Comment avez-vous fait pour visualiser le résultat de votre travail ?

5 Changement d'échelle

L'objectif de cette question est le même qu'à l'exercice 4. Toutefois, vous devez approcher le problème comme si on allait vous demander de pouvoir produire des 100aines de pages web du même style à partir de fichiers EAC-XML différents (bien que respectant la même structure). Comment procédez-vous ?

Pour valider votre travail, vous êtes invités à télécharger le fichier suivant et vous assurer que vous êtes parvenu à produire une page web pour chacun des fichiers contenus dans l'archive :

https://github.com/xgillard/lhist2532/raw/refs/heads/main/resources/exercice_5.zip

6 Butinage

Dans cet exercice, on complexifie encore un peu les choses : il vous est demandé d'aller butiner, moissonner, ou pour les anglophones faire du scraping, afin d'aller récupérer les images et transcriptions des documents disponibles ici : <https://www.oldbaileyonline.org/record/16770425>

Sur base de ces documents, vous créerez à chaque fois une page web html qui reprend le texte et l'image que vous avez récupéré. (Les plus aventureux d'entre vous peuvent ensuite essayer de modifier leur approche pour qu'elle permette de récupérer les images+textes des documents repris sur les trois pages suivante.)

Enfin, vous compléterez votre approche pour faire en sorte qu'outre les pages web liées à chacun des documents téléchargés automatiquement, votre approche soit capable de générer canvas de fichier EAD pré-rempli.

