

Bilan TER-3 10/10

ORDRE DU JOUR

- **Présentation de nos avancées**
- **Discussion sur les problèmes rencontrés**

DISCUSSION

- Comment extraire les questions/réponses et les intégrer dans la base de données
- Comment extraire les questions/réponses et les intégrer dans la base de données. Conversion des fichiers au format TXT via Pandoc (recommandé par le professeur).
- Notre problème concerne les questions/réponses : nous avons réussi à intégrer les lieux et d'autres éléments, mais pas les questions/réponses.

NOTES

- La version du document est en cours de réalisation par Sabrina, avec les nouvelles données communiquées par M. Collins bientôt.
- Pour les 4 lieux, il s'agit probablement d'un changement de nom : une réponse peut être associée à plusieurs noms. Il faudrait créer une table à part avec un auteur principal, et les liens externes seraient attribués via des IDs dans la base.
- Formater et extraire les données (SQLite) ; sinon, utiliser le format natif de Pandas (à vérifier).
- Les images ne sont pas prises en compte.
- En SQL : attribution des questions/réponses aux responsables.

- Pour l'instant, nous utilisons un fichier CSV comme base, puis l'injection se fera directement après sélection.
- L'exposé sur BERTopic est reporté à novembre.
- Spire.Doc a été testé, mais nous devons encore vérifier s'il convient à nos besoins.SQLAlchemy (convient pas).