

Compte rendu du 15/09/2023

Rencontre n°1 avec Sandra Bringay et Maximilien Servajean

SUJETS : Sport, INM et Covid

Etape n°1 du TER :

- Utiliser Tableau avec la licence étudiante
- Brainstorming sur les visualisations (sur papier) pour venir avec des propositions pour la prochaine fois
- Explorer les données avant et réfléchir à la problématique (Sources: API Twitter ou Reddit)
- Regarder sur Internet des visualisations
- Dessiner les maquettes des visualisations sur papier
- Avoir une discussion avec le doctorant pour avoir le même type de jeu de données
- Demander des conseils au doctorant sur les différentes visualisations et la méthode utilisée
- Déposer le compte-rendu sur Git-hub

Termes importants :

- Controverses = statistiques (exemples : combien de messages, algorithme de forces - graphes des utilisateurs avec les communautés qui se détachent pour les différentier)
- Bert Topic → mots clés importants, les thèmes
- TF-IDF : Statistique qui compte combien de fois un mot apparaît dans un poste multiplié par le log de l'inverse du nombre de fois qu'il apparaît dans tout le corpus
- INM = Interventions Non Medicamenteuses (jeu de données sur la Santé important pour Sandra)

Méthodes envisagées :

- Traitement des données à partir de base de données NoSQL
- Mettre en avant l'aspect temporel de la controverse

Prochaines dates de rencontres :

- 22/09 à 8h30
- 13/10 à 8h30
- 20/10
- 10/11
- 17/11 (seulement Maximilien Servajean)
- 8/12 : thèse du doctorant Samy
- 7/12
- Semaine du 11/12 : répétition de la soutenance
- 12/12 : présentation à Sandra Bringay
- 15/12