

## **Compte rendu – Semaine 3**

Au cours de cette semaine, nous avons réalisé les tâches suivantes :

- Nous avons ajouté des données dans nos deux bases. Bien qu'elles contenaient déjà un nombre suffisant d'exemples, nous avons constaté qu'il était possible d'en intégrer davantage afin notamment d'améliorer l'équilibrage.  
Pour ce fait : pour garantir un meilleur équilibre des données, nous avons limité la base de test à un maximum de 20 textes/lignes par catégorie de harcèlement, et fixé la base d'entraînement à 200 textes/lignes par catégorie.
- Nous avons créé notre base de données sous SQL en intégrant les deux bases (entraînement et test), et on a eu un seul fichier : harcelement.sql  
Nous avons également conçu le MCD (Modèle Conceptuel de Données) ainsi que le MOD (Modèle Organisationnel de Données).

## **Pour la semaine 4 :**

- Amélioration des données existantes et ajout d'une nouvelle catégorie regroupant les textes ne contenant aucun contenu haineux, afin d'enrichir et de mieux structurer la base de données.
- Conception et réalisation de modèles de visualisation des données, accompagnés d'explications détaillées pour faciliter l'analyse et l'interprétation des résultats.