

# **Manuel des Annotations**

## **Contexte:**

Nous sommes en train de créer de meilleurs outils pour la détection automatique des discours de haine en ligne. C'est dans ce but que nous avons compilé une base de données comportant quelques milliers de phrases courtes, chaque phrase reflétant un aspect particulier de l'expression de la haine en ligne.

## **Votre Mission:**

1. Notez chaque phrase que vous aurez à traiter comme **haineuse** ou **non haineuse**.
2. Signalez les phrases que vous jugez **irréalistes**.

## **Qu'entendons-nous par "haineux"?**

On définit comme **haineuse** une agression verbale dirigée vers un groupe protégé ou vers ses membres parce qu'ils font partie de ce groupe. Les groupes protégés sont basés sur l'âge, le handicap, la race (couleur de peau, nationalité, ethnicité ou pays d'origine), la religion ou les croyances, le sexe et l'orientation sexuelle aussi bien que l'identité sexuelle.

## **Qu'entendons-nous par "irréaliste"?**

Nous vous laissons intentionnellement le soin de le définir. Nous n'avons pas créé de phrases ayant pour objectif d'être irréalistes, donc l'option de signalement à l'aide d'un drapeau est uniquement mise à votre disposition afin d'assurer la qualité des données.

## **Ce à quoi il faut que vous fassiez attention pendant que vous annotez:**

- Dans certains contextes le langage haineux est employé de manière non haineuse (ex: une phrase de répartie qui fait référence à des insultes haineuses: "Ce n'est pas acceptable de traiter quelqu'un de nègre").
- La cible des abus est à prendre en compte pour juger que le contenu est haineux. Les abus interpersonnels et les abus dirigés contre un groupe non protégé comme ceux rattachés à une profession ou à une affiliation ne sont pas considérés comme haineux. (ex: "Je déteste les juifs" est haineux mais "Je te déteste", "Je déteste ma table" et "Je déteste les médecins" ne le sont pas).

## **Autres éléments clés:**

1. Évitez de noter des phrases comme étant irréalistes seulement parce que vous pensez qu'il est peu probable qu'elles apparaissent en ligne. Signalez uniquement les phrases incohérentes ou qui comportent des erreurs grammaticales.
2. Il vous est demandé d'annoter ces données de manière indépendante et d'éviter d'en parler autour de vous.
3. La plupart des données sont des phrases courtes, il est inutile de les analyser en profondeur.

**Merci infiniment pour votre travail!**