

2023

Using ChatGPT to generate Gendered Language

Shweta Soundararajan

Technological University Dublin, shweta.x.soundararajan@mytudublin.ie

Follow this and additional works at: <https://arrow.tudublin.ie/datas>



Part of the [Computer Sciences Commons](#)

Recommended Citation

Soundararajan, S. (2023). Using ChatGPT to generate Gendered Language [Data set]. Technological University Dublin. DOI: 10.21427/AYM7-SF29

This Dataset is brought to you for free and open access by the Data@TU Dublin at ARROW@TU Dublin. It has been accepted for inclusion in Datasets by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie, vera.kilshaw@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 International License](#).
Funder: Technological University Dublin

Datasets for Gendered Language generated using ChatGPT

Files description:

Three datasets(set 1, set 2, set 3) generated using Gaucher lexicon are:

1. Gaucher_dataset_set_1.csv
2. Gaucher_dataset_set_2.csv
3. Gaucher_dataset_set_3.csv

Three datasets(set 1, set 2, set 3) generated using Cryan lexicon are:

1. Cryan_dataset_set_1.csv
2. Cryan_dataset_set_2.csv
3. Cryan_dataset_set_3.csv

Each CSV file has the following columns:

1. Index: Index of the sentence.
2. Sentences: List of sentences.
3. Labels: Labels assigned to the sentences. The labels are 'consistent' and 'contradictory'.
4. about_Male: Has two values, either 0 or 1. 1 indicates that the sentence is about male, and 0 indicates it's about female.
5. Masc_terms: Has two values, either 0 or 1. 1 indicates that the sentence is generated using male-coded words from the lexicon, and 0 indicates it's generated using female-coded words from the lexicon.