Who's That Author? Authorship attribution of free online texts with computational stylometry.

Abhinil Dutt*
Andrew McQueen*
Murray Ahmed*
University of Illinois Urbana-Champaign
Champaign, IL, USA

1 SUMMARY

2 INTRO

3 DESCRIPTION

3.1 Data Collection

A web-crawler was written to traverse Project Gutenberg and extract metadata about each author and their written works.

3.2 Data Exploration

4 INITIAL EVALUATION

5 INSIGHTS

All data retreived from Project Gutenberg comes in a very raw and unprocessed format which requires substantial preprocessing.

---

*All authors contributed equally to this project.