**Faculty Name:**

Kaizheng Wang

**Faculty Email:**

kw2934@columbia.edu

**Lab:**

Wang lab

**Project Title:**

Integrative analysis of text corpora

**Description:**

This project aims to develop methods for simultaneous analysis of multiple text corpora. The student and the mentor will (1) build statistical models for studying the commonalities and differences among different corpora; (2) develop efficient inferential tools; (3) collect and clean real-world corpora to test the proposed methods; (4) conduct a mathematical analysis to establish theoretical guarantees; (5) write a paper to summarize the results. The student needs to (1) have past experience demonstrating independent research capabilities; (2) have taken courses or conducted research on statistics, machine learning, optimization and real analysis; (3) be proficient in Python and LaTeX.

**Location of Research:**

Hybrid (both remote and on-site)

**# of hrs/week:**

40

**Department/Program:**

Industrial Engineering & Operations Research

**Eligibility:**

BS, Second Year, BS, Third Year

**To apply, please contact:**

Kaizheng Wang, kw2934@columbia.edu