

# Data Science Project

*Gege Gui*

*9/11/2017*

## Project Overview:

Perform an analysis of the statistical analyses in all published PLoS papers. What are the most common techniques? How do they vary by field? Are there any trends over the last 10-15 years?

- Follow the documentation from PLOS API website <http://api.plos.org/solr/examples/> to download statistics-related paper.
- Change pdf files to txt files, extract key words and publishing time from each paper.
- List common statistics techniques using several textbooks. (or from all key words)
- Read paper: “Topics over Time: A Non-Markov Continuous-Time Model of Topical Trends” and try to extend the LDA model in this project