

**Group: 10****Groupmembers:**

Alex Lopez #026664623

Anthony Tran #030061354

Glen Lee #027083106

**Topic Choice: Cloud-Based Web Scraping****Inspiration:**

- <https://ieeexplore.ieee.org/document/8118431> [1]
- <https://ieeexplore.ieee.org/document/10270597> [2]

**Approach:**

Our choice for this project is to work on a web scraper that is built on a distributed system with the use of Google Cloud. Inspired by a research paper[1] using AWS EC2 instances and Selenium to create their own distributed web scraper, our team would like to try and create this on the Google Cloud Platform as opposed to Amazon's AWS. The aforementioned research paper focuses their software on data mining for Big data applications. We would like to apply this data mining for food recipe aggregation[2]. Our goal is to show that if our software is able to fill a database such as a food recipe database, then there are many types of databases that can also be populated.