

Team 2: Edona Mujaj [edona@bu.edu](mailto:edona@bu.edu), Andrew Choi [dooby@bu.edu](mailto:dooby@bu.edu), Colette Li [ruojiali@bu.edu](mailto:ruojiali@bu.edu), Sangjoon Lee, [sj0726@bu.edu](mailto:sj0726@bu.edu)

- What is the project focus/overall goal?
  - Our project has two main goals, the first is to identify the difference in gender analysis when analyzing characters vs US translated names as well as compare the gender makeup of inventors on Chinese patents and Korean patents.
- What type of data will you collect or be analyzing?
  - The data we will analyze are the patents that were submitted in the US and in China or Korea.
- What are potential limitations of the project?
  - A potential limitation is the possibility of a language barrier, not all of us are fluent in both, which could lead to difficulties in analyzing the data. Another limitation could be that the names in Chinese and Korean do not offer any help with determining the gender of the inventor.
- Why is this project important or Why did you choose this project?
  - This project is important because it can help institutions distinguish and predict gender based off of their name. It can save a lot of time, and help quality check data. We chose this project because we are interested in Chinese and Korean culture, especially in the linguistics aspect. It will be intriguing to see if there are any patterns or correlations between names.
- What is your team's next steps? (include action items/tasks)
  - Our next step is to look over the data and work on character recognition for Chinese and Korean Characters.