Connecting CMU Researchers to Current Grant Opportunities
By: Divya Rao, Pam Sun, Lisa Zhang
Advisor: Peter Freeman

Aim
- In Fall 2020, a group of CMU Statistics & Data Science undergraduates developed a data exploration tool to explore and analyze data on grant funding involving CMU researchers. As an extension to this work, we are interested in building a data recommendation system that matches previously funded CMU researchers (as documented in one database) to future grant solicitations (as provided in another).
- We utilize the Dimensions (previously funded) and GrantForward (future solicitations) databases.

Data
- The data were provided by Neelam Bharti, Huajin Wang, and Sarah Young of the CMU Libraries.
- The Dimensions dataset is a subset of 6,326 previously funded research grants affiliated with CMU from the Dimensions AI database. This dataset consists of information about the grants such as title, abstract, funding amount, start/end date, researchers, funder/funding group, resulting publication, grant URL, and categorizations. We leverage the data from six categorization systems to generate the subject list we use when matching across datasets.
- The GrantForward dataset is a subset of 2,856 currently available research grant opportunities with solicitations from September to October 2021. This dataset consists of information such as title, description, sponsors, funding per grant, grant URL, and subject.

Matching Opportunities to Researchers
- Use subject column to obtain specific topics, and more generalized ones with fuzzy matching
- Create a list of related researchers through the matching of obtained topics to category columns in dataset
- For each related researcher, quantify the correlation between the words in the grant description, publication description, and researcher abstract
- Rank the list of related researchers based on word-wise correlation and return the top ten matching researchers, along with related information

Future Work
- As mentioned above, our recommendation system currently uses a subset of data from the GrantForward database. In the future, we aim to tie our recommendation system to the GrantForward database by utilizing its API.

Grant Opportunities by Category
List of relevant main and subcategories

References
- Dimensions, https://www.dimensions.ai/
- The Fall 2020 Project: https://www.youtube.com/watch?v=ijZXBx82ExA