

Survey Search

...

by Pawsitive

CS 30122 Project

<https://github.com/bethbailey/survey-search>

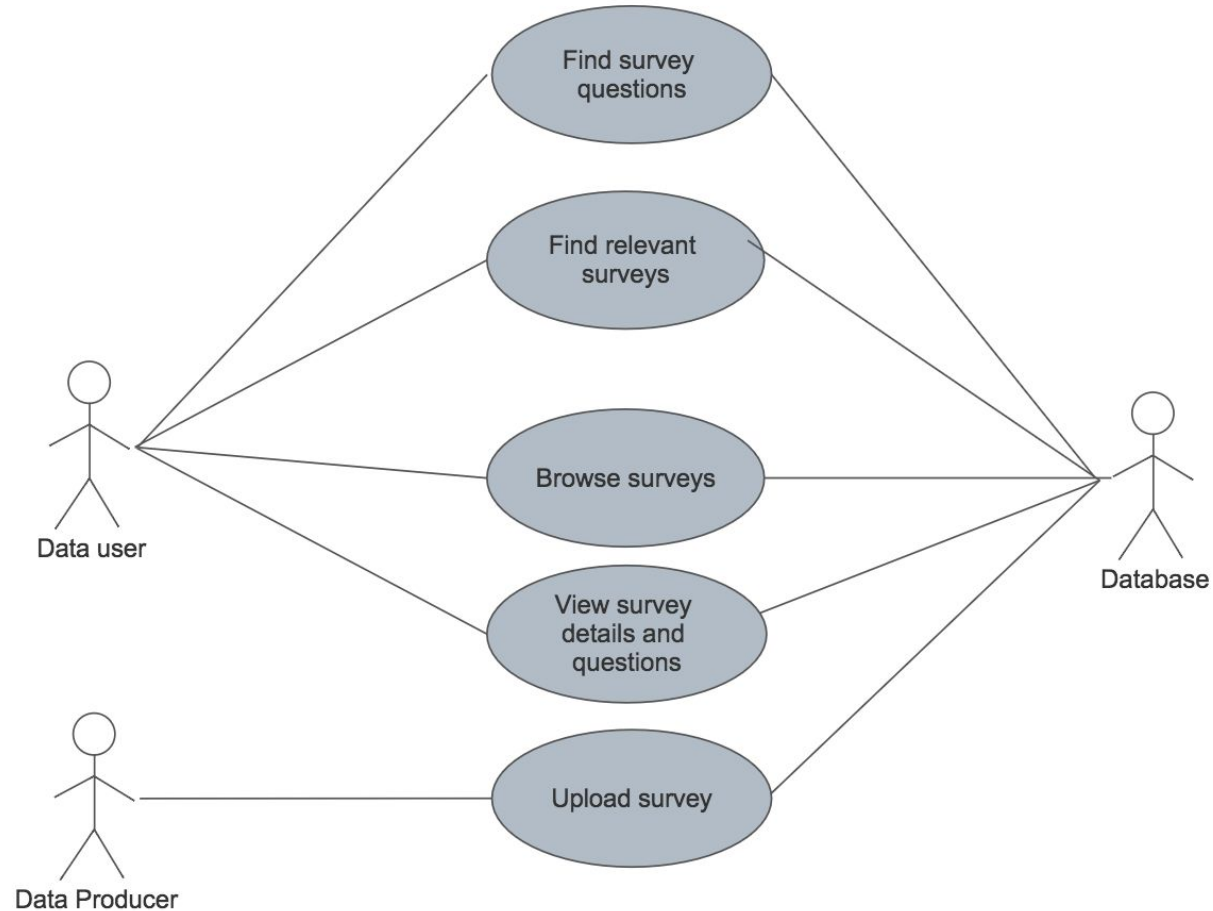
Project Overview

GOAL: Create an easy-to-use survey data repository.

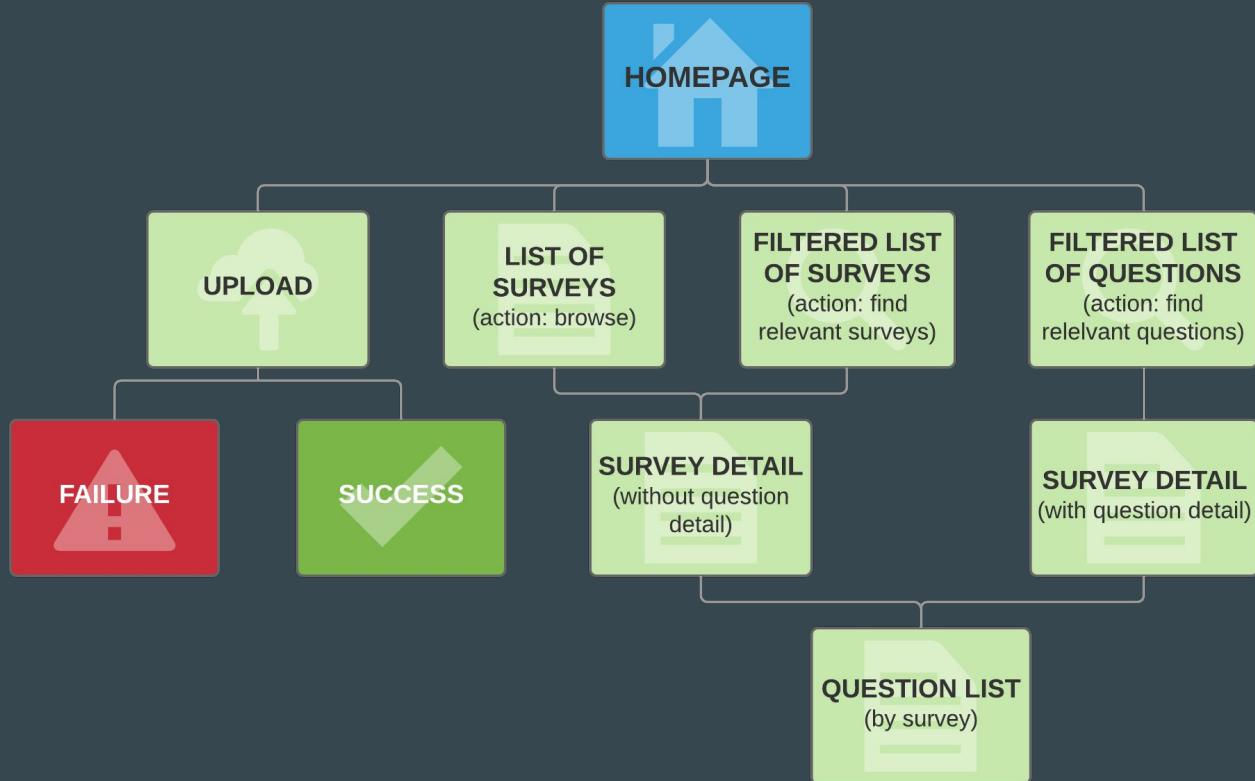
What makes this **different** than other data repositories? Users can:

1. Search *multiple surveys* from different sources.
2. Search surveys on a *granular level* (for individual questions) as well as for entire surveys
3. *Upload and share* their own survey data

Use Case Diagram: Survey Search

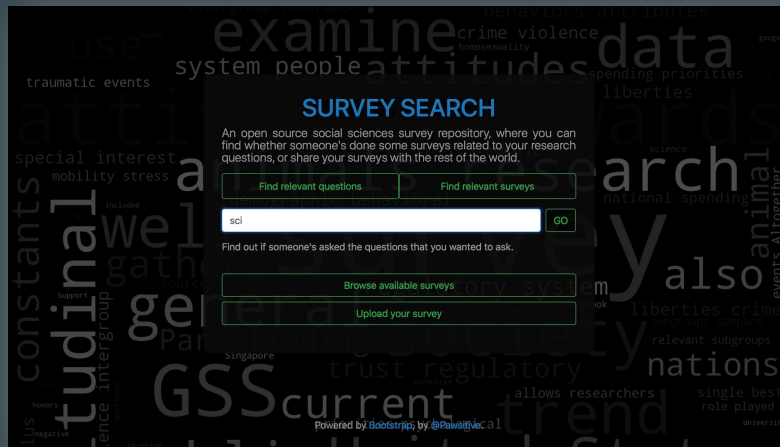


PAWSITIVE SITE MAP



Cool Stuff We Learned!

- Data processing/handling files in Python (pypdf, docx2txt)
- Lucidchart
- Building a website with Django
 - Implement TF-IDF based custom search algorithm
 - Bootstrap frontend
- Word cloud generation



Challenges

- Survey data cleaning is HARD
 - Data is in many formats
 - Difficult to reuse code
- Django's modularity makes it versatile, but learning curve is steep
 - Integrating multiple modules: search, form, files, word cloud, etc.
- Learning the best practices of software development
 - Version control
 - Teamwork is hard, communicate with and love your teammates <3
 - A lot of programming involves using resources and reading documentation
 - Don't be afraid to try things

Features to come?

- Shared search history/autocomplete
 - Visualization of search patterns
- Advanced query engine
 - A smarter search that accounts for typos and similar words
- Visualizations of survey data

Let's do a demo!

Questions/Comments?

