

Team 096 Project Proposal

Team Name: OpenSource

Project Summary:

Our project aims to connect data in the academic world to help aspiring PHD candidates select a university and mentor that best fits their interests. Our projects will parse the research output (papers, studies, experiments, etc) of a number of universities, define a set of topics, and identify a number of experts in those fields. During account creation, users will be able to select the topics they are interested in and we will tune their search output by that. Users can search for specific experts/mentors, and universities. We will output a graph that shows the relationship between different experts in their selected fields and how that is tied to universities. Nodes in this graph will be Experts and Universities while edges will indicate that there can be some collaborative research output by two experts and/or is from a university. Users can select a node to see more information on that person or university.

Description

Much like how many students felt looking for a college major or university to attend, PHD candidates or aspiring graduate students are overwhelmed with choices. They may not know what the collaborative atmosphere looks like for a particular topic or what the leading experts or universities. This tool aims to help alleviate that. Even for seasoned experts or hobbyists, this tool will help give them a gauge to how a field is doing by glancing at an easy to read graph.

The flagship page for this website will be the graph we populate for each topic. We will have nodes for all experts and universities with edges representing some sort of collaboration between nodes. For example, if an ECE student considering academia wants to see how collaborative the quantum space is, they can enter that into the graph populator and see who is making progress in that field and how collaborative they are. They'll get a sense for the best universities in quantum and who some of the leading experts are. If the students really enjoyed group projects in their undergrad and are looking for a very collaborative field and see that nodes in the quantum topic graph have a large number of edges, they may see that this is a good fit. If they want to move forward with their decision, they can use our suggested experts and universities feature to see the top universities and researchers in that field that they may want to apply to or contact. Users can also build new boards (while saving old boards) with a new set of topics.

Moreover, while a user has selected an edge and is scanning through information on some research output, we will select some music to play in the background. The music genre will be selected based on the topic of the research paper they are reading (we will arbitrarily decide that, for example, math papers pair with jazz music). The actual song within a genre they get may be random so users don't get bored when reusing the same board.

Technically Challenging Features:

We will use some tools to help us visualize data. For example, the [Network Chart from Python Graph Gallery](#) would transform the relations between objects (Experts and Universities) that come out from the database into a visualized graph and it is easier for people to understand. Edges in this graph will represent some sort of research output. An edge between two experts indicates they published a paper together or had some other sort of collaborative effort. An edge from an expert to a university indicates that person is affiliated with that university. By selecting on an object/node, a user can see more information on that expert or university. The same is true for an edge. Moreover, we will implement an algorithm that determines what universities and experts are best in a particular field.

Usefulness:

This project can be useful for students looking to go into academia. If they are interested in a particular field/topic, they can input that and see the sort of people who are experts in that field, the universities that boast the strongest programs, and the sort of collaboration they can expect. This tool can also help potential PHD candidates in their search for a mentor and university. If in their undergrad they really enjoyed working with a particular professor, they can see who else collaborated with that professor and what topics they are top experts in. Also, the music feature will keep the user engaged while they go through what may otherwise be tedious research.

Realness:

Dataset 1 (To get different research outputs):

<https://docs.openalex.org/api-entities/works>

Dataset 2 (To get information on experts):

<https://api.openalex.org/authors>

Dataset 3 (For which institutions publish works):

<https://api.openalex.org/sources> and <https://docs.openalex.org/api-entities/institutions>

Dataset 4 (To categorize topics):

<https://docs.openalex.org/api-entities/topics/get-lists-of-topics>

Dataset 5 (To get research output and collaboration):

<https://www.kaggle.com/datasets/mathurinache/citation-network-dataset>

Dataset 6 (For music files categorized by genre)

<https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification-data>

Dataset 7: (We will scrape LinkedIn for information on where experts got their PHD)

lix-it.com, or similar service. We can also write our own script if existing ones are insufficient.

Functionality:

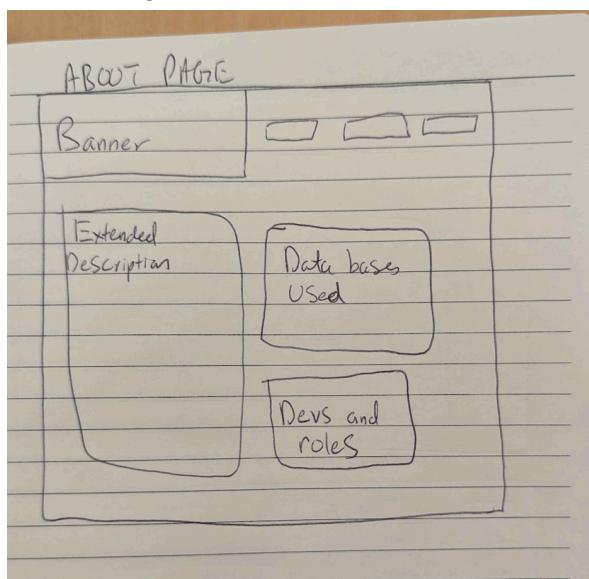
Users of the website will be able to do the following:

1. Create or log into their account
2. Select from a list of topics what they are interested in
3. Search for an expert or university as a starting point
4. See a graph populate with the following
 - a. Nodes for experts or universities
 - b. Edges for an affiliation to a university or collaborative research paper/output
5. See additional on any selected node or edge
 - a. Listen to music while reading through details on a paper
6. Build new boards with new topics/fields while saving old boards.
7. See a list of suggested experts or universities based on the user's interests.

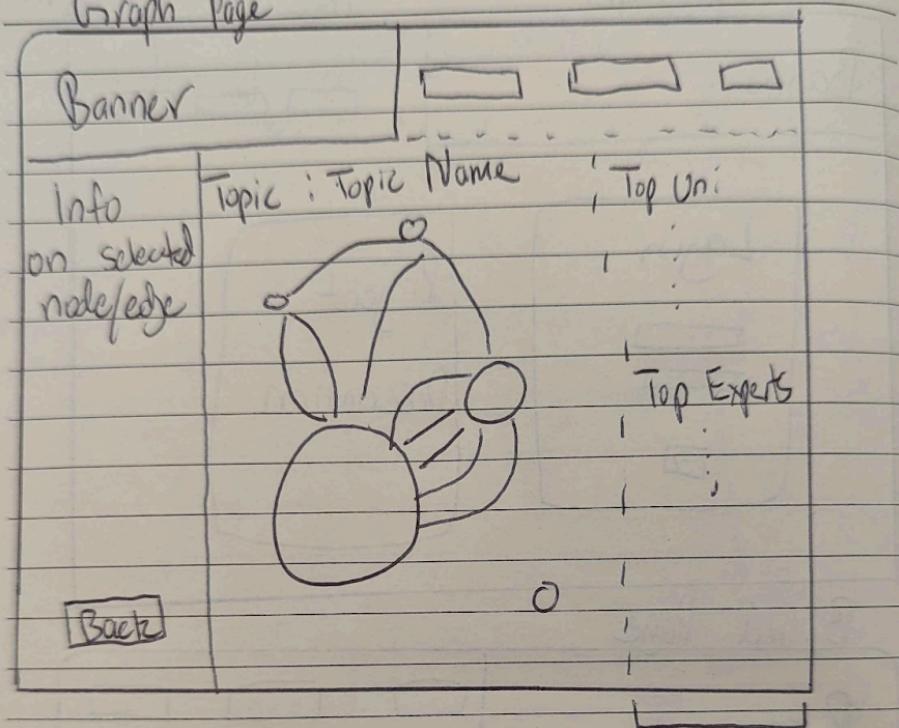
Work Distribution:

- a. Recommendation Algorithm for relationship suggestion: Alex
- b. Frontend Page layout/infrastructure: Philip James
- c. Frontend Flowchart/board visual: Chinmay
- d. User authentication: Philip
- e. Backend API Design: Boyang
- f. API implementation: All
- g. Saving old boards: Chinmay
- h. Creation and querying of our databases: Philip, Alex
- i. Connecting to and querying external databases: Boyang, Chinmay
- j. Suggestion algorithm: Boyang, Alex
- k. Assign each paper a music genre and randomized song selection: Philip

UI Mockup:

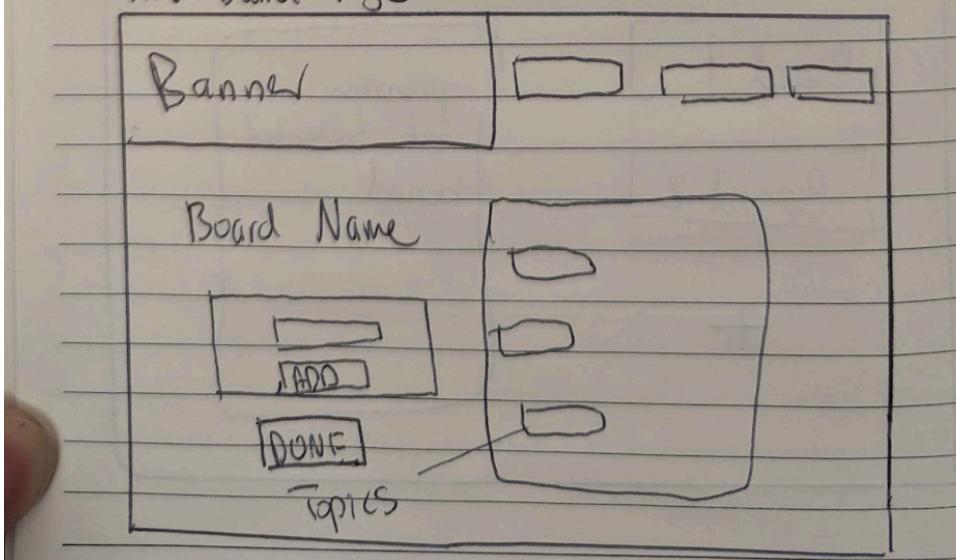


Graph Page



New Board Page

Side bar pop out



Welcome Page

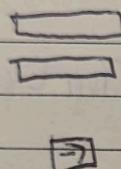
Board
Page

About

Login

Banner

Login



Project

Description

Board Page

Banner

Board 1

Board 2

+

preview
of selected
board