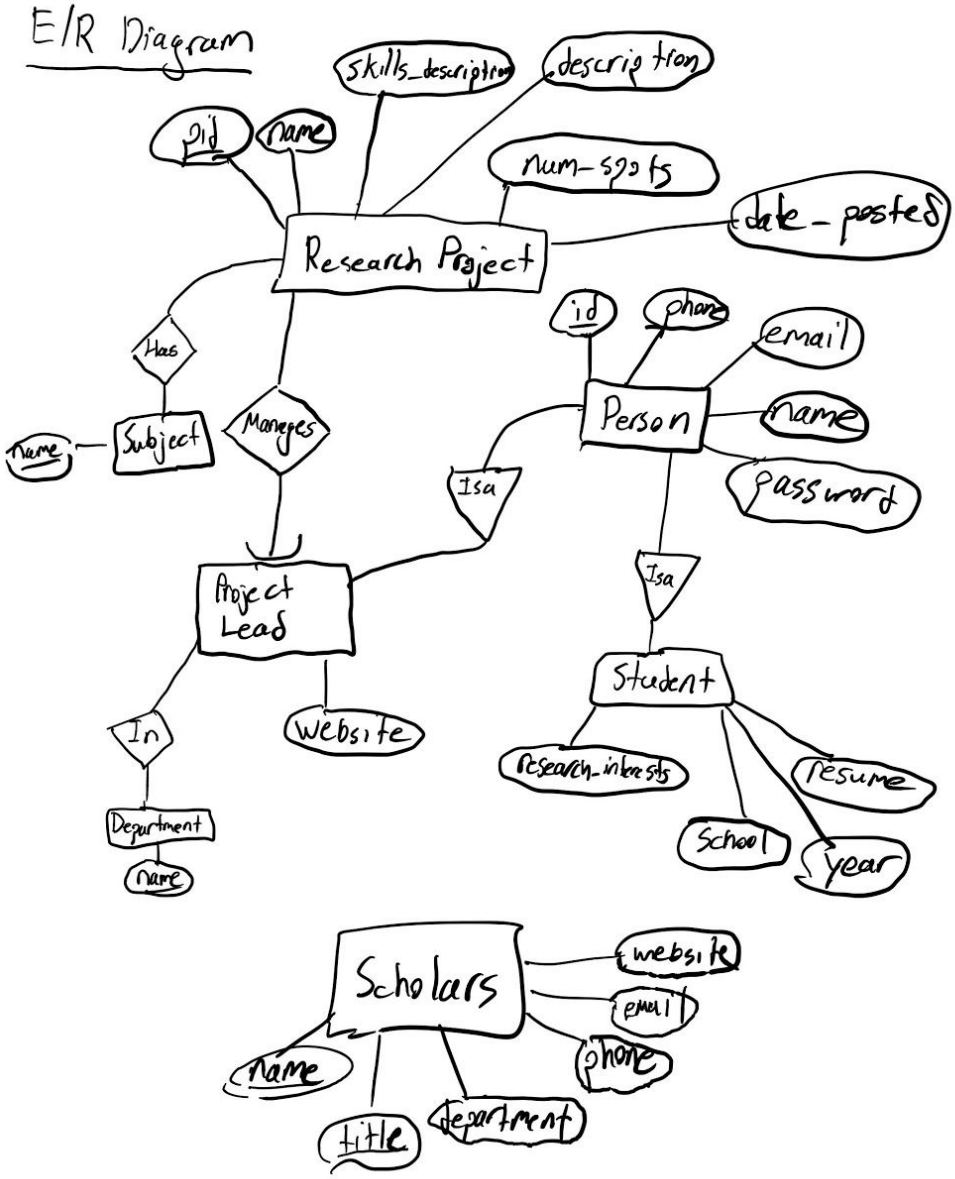
“Duke ResearchHub” Final Project Report

Christian Burke, Samuel Chan, Vinit Parekh, James Rumsey, Justin Suh
CompSci 316: Introduction to Databases
12/11/2018

A brief description of our application

Our group is building a webapp that will allow Duke faculty to upload postings for research projects at Duke. Our goal is to connect students to research opportunities by providing them up to date information that they need in order to determine which lab to research in. We will be developing a query system that would allow students to look for projects that have been posted, or look for Duke faculty to directly contact. Each project page will have information about what work the project entails, the number of spots available, a brief description of the necessary skills, and the date posted. Project Managers have the ability to register on the website and post projects for students to public brows. Students can search for a project on the project page and apply through an apply button that redirects to a form to fill out. Alternatively ,Students can search a separate faculty page that displays the name, title, department, phone number, email, and website of each faculty member at Duke. Information on faculty was scraped from Duke’s faculty directory (scholars.duke.edu).

E/R Diagram


E/R Diagram
skills_description
description
pid
name
num_spots
date_posted
Research Project
Has
Subject
name
Manages
Project Lead
In
Department
name
website
id
phone
email
Person
name
password
Isa
Isa
Student
research_interests
resume
School
year
Scholars
website
email
phone
name
title
department

Assumptions

We implemented a Project page, which is a dynamic page of projects and their details, and the Scholars page, which will be a static list of the faculty members associated with research. The current project data is randomly generated. Before randomly generating the data, we performed interviews with students to understand what kind of data they would like presented to them when searching for projects. The Scholars data is web scraped from https://scholars.duke.edu/. Based on the interviews with students and the data available on the website, we gathered each faculty member's name, website, etc. to display under the Scholars section (Scholars table in the PostgreSQL database). We hope that students can use this page to obtain information about professors and email them, even if they do not have explicit projects posted. We gathered data to fill the database, at least initially, by randomly generating fields for the Project tables and web scraping for the faculty tables. When it came to practically designing the website layout, we realized that it would be extremely impractical to implement many of the features we had planned at the beginning, so we were forced to prioritize what we believed to be most important. Because of this, we decided to scrap the Student tables and page, but we still believe that it could be a useful feature that could be implemented if we decide to continue with this project.

List of Tables

Person(id, name, email, password, phone, admin)

- The Person table holds the unique id for each user, their name, email, password, and phone number. The admin column is boolean and can be used in further versions of the website. Admins will have special permissions (e.g. deleting projects and user profiles, edit access) .

LeadInfo(id, website)

- This table references rows in Person. Users that are managing projects are added to this table. The website column is also here because faculty may want to include a research group/lab website on their profile while a student may not.

StudentInfo(id, research_interests, resume, school ,year)

- Like LeadInfo, this table references rows in Person. This table is unused in our front-end, but in the future, we hope that students will be able to create public profiles similar to LinkedIn. PIs will be able to browse these to view potential candidates.

ProjectInfo(pid, name, num_spots, date_posted, description, skills_description)

- The ProjectInfo table holds information about each individual project.

ProjectSubject(pid, subject)

- The ProjectSubject table holds the subjects that each projects encompasses. This is a separate table because a project may encompass multiple subjects (e.g.. a project on machine learning may interest students with interests in Computer Science or Mathematics).

Scholars(name, title, department, phone, email, website)

- The Scholars table holds the information that was gathered from our web scraping application that gathered the data of Duke faculty.