

[COSC-257] Design Document: Being Human in STEM Database

Authors: Hanaa Charania, Nayeon Shin, Rachel Lin, and Hewan Worku

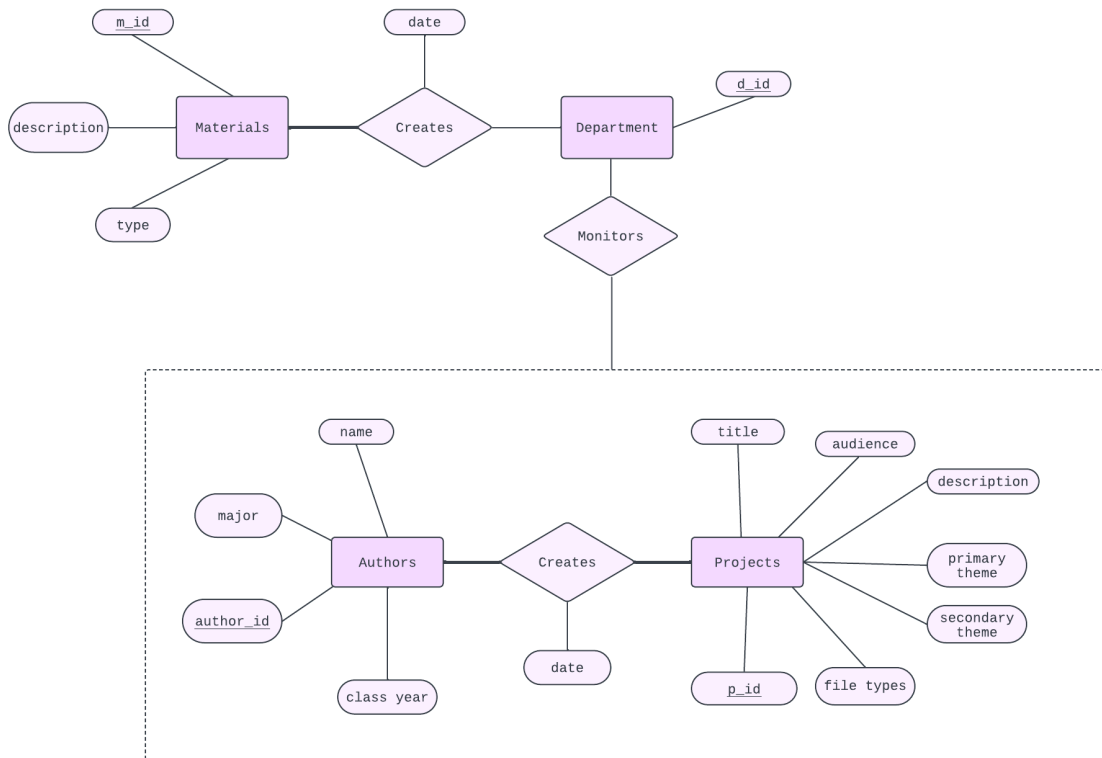
Date: September 23, 2023

Description

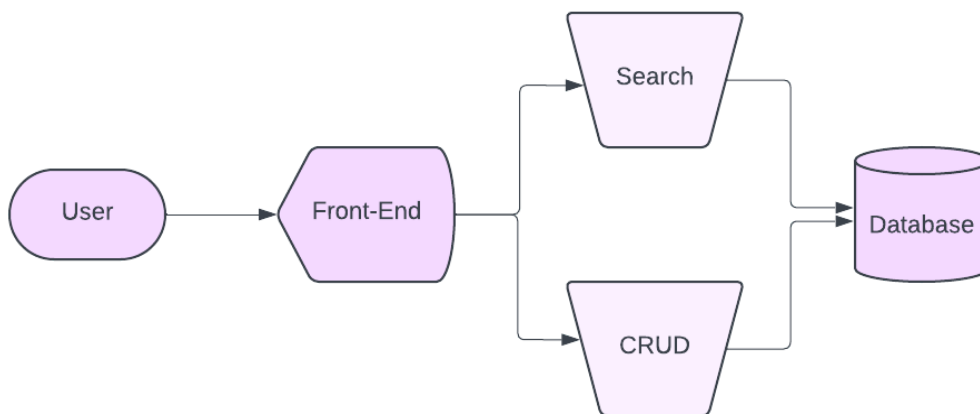
The Being Human in STEM (HSTEM) initiative seeks to foster a more inclusive STEM community where students of all backgrounds feel comfortable speaking about their individual identities and diverse experiences in academia. HSTEM empowers students from a variety of universities to investigate issues of racism and equity in STEM through community engagement on their campuses. HSTEM started off a special topics course at Amherst College in the Spring of 2016 after the Amherst Uprising. Since then, it has evolved into a student-driven initiative with projects on DEI, dissecting academia, humanism and wellbeing, and inclusive outreach. Other HSTEM materials include presentations, documentation on how to throw a public salon, archival records, a documentary, etc.

With our Being Human in STEM Database project, we aim to organize these materials in a more accessible manner where users can search for specific information from talks, recordings, transcripts, etc. Files will be searchable based on tags: author, iteration (Spring 2016, Fall 2020, etc.), major, theme (DEI, dissecting academia, etc.), and audience (faculty, international, etc.) This database is directed at Amherst students and faculty/staff, and beyond. It will be similar to a “living” hub in that Prof. Jaswal (our client and the founder of HSTEM) can continue adding materials and additional tags to it as the course grows each year. To gather our data, we will be “scraping” the existing HSTEM website (<https://www.beinghumaninstem.com>) and utilizing the current HSTEM Google Drive. Our final product will be a web application that we will deliver to our HSTEM clients.

ER Diagram



Interactions with Database



Our database will serve as a dynamic, living repository and have a user-facing front-end. Separating the user interfaces from the database will serve as an abstract layer, eliminating the need for users to learn how to manipulate the database directly. The UI will enable the users to search and perform CRUD (create, read, update, delete) operations. Users can also search for resources using identifiers such as title, discipline, format, and audience. Additionally, they can search for specific titles like talks, recordings, and transcripts.

Current State of Data

The current state of the data is organized in Google Drive with a folder system containing all of the files according to a specific category. The majority of the folders each contain previous project files or video recordings based on a specific iteration of HSTEM (e.g. year and semester). The types of files for each project include PDFs, JPG images, google slide files, google document files, excel spreadsheets, and MP4 video recordings. Professor Jaswal is currently keeping track of all the HSTEM projects for specific years (e.g. as of Spring 2023 or 2017) using a Google Sheets spreadsheet that lists the project number, name, and link to the respective project folder or document (e.g. PDF). Similarly, there is a master document that lists all Amherst HSTEM projects that contains the iteration, student name, major, class year, project title, project link, link to summary slide, and a checkbox corresponding to the project topic/tag, followed by comments about the project. Most of the project tags/comments are not checked off or filled in, so it may be difficult to keep track of descriptions for each project. Additionally, the links on the spreadsheet for each project folder or document may make it challenging for easy access to the data, particularly on the user side.

Software & Deliverables

We are planning to deliver a web application to our clients at HTSEM with ReactJS on our front-end and Express and PostgreSQL on our back-end. We want to host our website on the cloud using AWS (Amazon Web Services). On the front end, the user can search for files or projects using tags (title, audience, author, iteration, file type(s), etc.), and the database will gather and display results based on keyword matches. The resulting files will be easily accessible and downloadable. Our client will be able to add further tags, descriptions, and items to the database via the web application after user authentication. Given the time constraints, our database may lack scalability and contain a low-complexity structure. There might be difficulties when modifying or linking every single type of file to the database from the given data on Google Drive and the HSTEM website. However, these limitations can be addressed in the future.