Capstone Project Final Report

Democratizing the Access to Opportunities for Career Exploration

May 10th, 2022

Lisa Guiotoko
Preeti Kumari
Ruoyu Lu
Lidia Ortiz Zamora
Sanjana Parmar
Velie-Velia Sando
Tina Wu
Zimo Yang

ACKNOWLEDGEMENTS

We would like to acknowledge with much appreciation the crucial role of the leadership team at HCV: **Anton Reed**, Director of Software and **Walter Lewis**, President & CEO.
We would also like to express our special gratitude to our advisor, Professor Lee Branstetter for his valuable guidance and support for the project. And Andrew Breazeale for his direction and enthusiasm.
Last but not least, we would like to express our gratitude to the following people for sharing valuable insights with us on our capstone project within Carnegie Mellon University's Heinz College of Information Systems and Public Policy:
Dan LaVallee, Senior Director at UPMC
David Mosey, Executive Director of Smart Futures
Dr. Judy Hallinen, Ph.D. , Director, Leonard Gelfand Center for Service Learning and Outreach
Dr. Justin Agilo, Ph.D., Executive Director of Readiness Institute, PennState
Dr. Lisa Abel-Palmieri, Ph.D., President and CEO, Boys & Girls Club of Western Pennsylvania
Dr. Robert Scherrer, Ph.D., Executive Director, Allegheny Intermediate Unit
Geng Wang, CEO and Founder, Civic Champs
Jace Ransom, School Age Youth Manager at DHS
Jackie Foor, COO at the Consortium for Public Education
Mayada Christiansen., Director of Youth Workforce Programs, Partner4Work
Michelle Howard, Coordinator at STEMisphere
Richard LaPratt, Database Director at 211 United
Sharmyn Straughters, Director at CCAC

Table of Contents

Executive Summary 3

Project Objectives 4

Analytic Approach 5

Research **5**

Database 5

Web Application 6

Data Collection **7**

Database Creation **8**

Entity-Relationship Model 8

Normalization 9

Database Creation 10

Web App Development **10**

Front-end Development 10

Back-end Development 11

Data Analysis **13**

Data Summary **13**

Geographical Analysis **15**

Opportunity in the Spotlight **17**

Conclusion **18**

Recommendations 18

Shift to contributor-based model **19**

Database Maintenance **21**

Design a Success Matrix **21**

Awareness Program **22**

Logistics and Safety **22**

References **24**

Appendices 25

Appendix I - Outreach Process **25**

Appendix II - Database Documents **26**

Data Dictionary 26

Entity-Relationship Diagram 29

Script 30

Appendix III - Web App **32**

Appendix IV - Distribution of Opportunities **33**

Appendix V - Map **34**

Executive Summary

Children growing up in Pittsburgh's lower-income families often lack a direct connection to the knowledge-intensive industries that are remaking the city's economy and our nation in the 21st-century. A lack of exposure to the opportunities presented by these emerging industries, a lack of understanding of the skills they require, and a shortage of mentors who can guide students into these careers prevent many of Pittsburgh's children from fully sharing in the prosperity these industries are bringing to their employees and shareholders. Nonprofits, universities, museums, foundations, and corporations recognize this barrier, but their efforts to contend with it are often uncoordinated and fail to reach the children they seek to benefit. As a solution to the issue described above, the team collected data on local opportunities relating to career exposure, exploration, and experience within Pittsburgh. We also created a database of these opportunities and developed key components of a web application that will help connect HCV-supported students to these opportunities.

Project Objectives

A 2017 study by Alex Bell and colleagues on Who Becomes an Inventor in America, shows that children with parents in the top 1% of the income distribution are ten times more likely to become inventors than children with below-median income-earning parents.[1] Additionally, an experiment they studied on a Big Brothers Big Sisters mentorship program proved that the mentorship program yielded positive behavioral and social outcomes for disadvantaged youth.[2]

These studies make clear that lack of exposure to opportunities, limited access to industries, limited opportunities to build skills that would enable success in the workforce, and shortages of mentors yield unfavorable outcomes for disadvantaged youth.

Having identified the issue, HCV is working on a project called SMART Dreams for students to explore career opportunities, manage progress through high school, and plan for life after graduation. In partnership with HCV, this project aimed to the following:

1. Collect Career Education/Exposure/Experience (CE^3) for 21st-century Pittsburgh middle and high school students
2. Create a database using MySQL of the Career Education/Exposure/Experience (CE^3) opportunities
3. Create a web-based app that can help match interested students supported by HCV with appropriate experiences, internships, and mentorships that can help connect them to 21st-century career opportunities.

To achieve the goals highlighted above, we aimed to identify organizations that would have pertinent opportunities, locate and gather data from these opportunities, build a database, and develop key components of a web application for HCV to utilize. The database and web application we created to fit into this bigger picture as a tool to help connect students to opportunities and successfully plan for life post-graduation.

[1] The Quarterly Journal of Economics, Volume 134, Issue 2, May 2019, Pages 647–713, https://doi.org/10.1093/qje/qjy028

[2] Bell, Alexander. 2020. Essays on Income Inequality. Doctoral Dissertation, Harvard University, Graduate School of Arts and Sciences.

Analytic Approach

Aiming to create an up-to-date database with opportunities for Pittsburgh middle and high school students, the systems project team was divided into two sub-teams: the Database team and the App team.

The database team worked with Homewood Children's Village(HCV) leaders, and staff to obtain information about the needs of HCV students and their families for CE^3 experiences and services. The team also collaborated with CMU's Gelfand Center, the Carnegie Museums, the Children's Museum, MCG, and other organizations with experience and investment in providing CE^3 experiences and services to Pittsburgh youth to build out as comprehensive a menu of CE^3 programs as possible, identifying "holes" in the network of existing experiences, internships, and mentorships that local foundations and other funders might seek to fill.

The app team corporated with HCV's in-house developers to create components of a user-friendly app that can help students and their parents connect their students' interests to existing opportunities.

Research

Database

Our systems team was assigned advisory board members who were experienced in this field of work to answer questions and provide resources in our efforts to create a database of opportunities for the youth. We interviewed over 20 members and additional stakeholders, who gave us multiple insights about database creation, maintenance, outreach, etc. It became clear that there were patterns that most organizations faced when looking to create/maintain a database, and these outlines are outlined below–

Database maintenance: An important finding that was prevalent across the board was that creating and maintaining a database was quite difficult. From our conversations, we concluded that the hardships with database maintenance are a result of 2 main factors: shortages in human capital, and lack of connections within and between organizations. Many program organizers were very transparent and made it clear that unfortunately, their organizations found it difficult to frequently update the opportunities displayed on their own websites, often due to shortages in staff as their points of contact for their opportunities sometimes changed jobs. This brings us to the second component that plays into hardships in database maintenance, the lack of connections. The lack of connections within and between the organization was a result of one or two points of contact assigned to the database maintenance process. This posed an issue, especially when that point of contact left the organization because other staff members had not built the relationships necessary for the maintenance process.

Program outreach: There are external factors that hindered youth to access their desired programs, which organizations need to be aware of although there may not be much the organization can do to resolve this issue. The first issue comes from the information gap. Most programs relied on schools to provide the information to students and families, which did not always work in the organization's favor. Second, transportation often poses a problem for students to get to and from the programs, as these opportunities are often concentrated in one part of town. Third, the space limitation also restricts the programs' capability to hold a certain number of students. Especially for those free in-person activities, the demand is always greater than the supply.

Collaborations: There are a plethora of organizations in the community with similar goals to serve underrepresented youth. Many organizations have found that there is a lack of collaboration between these organizations, which often leads to issues with outreach and database maintenance. Collaborative efforts would facilitate outreach efforts to the target population, enhance database creation and maintenance, and enlarge the organizations' networks.

Organizational financial structures: As organizational financial structures dictate the types of funds they receive, it behooves the organization to make their financial structure conducive for receiving funds. This is because eligibility requirements, especially at the state and federal level, are very specific so organizations like Partner4Work that acquire these funds and are charged with disbursement at the local level utilize tools like Request for Proposals and the Pennsylvania Department of Education requirements to determine an organization's eligibility to receive funds.

Program scope and services: Upon conversing with Geng Wang, the founder of Civic Champs, there seemed to be certain steps to take to ensure the success of a preliminary web application. Mr.Wang suggested that the focus of the web application should be on features that have the highest value-add, which in this case is connecting underserved youth to opportunities. Focusing on the most important aspect of the application narrows the scope of the program services that HCVis looking to provide for viewers and drives the purpose of the application altogether.

Web Application

Before we started creating the Web App, we met with Geng Wang from Civic Champs to obtain insights into the process. Geng had rich experience in building products from scratch and had a series of successful entrepreneurial attempts. During the interview, Geng shared his experience of developing the Mentor App (how the idea came out and the challenges they have faced). He also provided suggestions for our App design work. Here are the highlights.

The development of an App requires dedicated human labor input and continuous upgrading and maintenance work. Take the Mentor App as an example, it took Geng's team 3-4 months to finish the MVP version with a development team of 2-to 3 full-time experienced engineers. After the MVP, it took them another 1-2 months to fix bugs.

Geng suggested we focus on specific features that have the highest value-add at the beginning. For example, Civic Champs started as a checking App that contains only check-in and check-out features. Geng also suggested that we should arrange enough time to obtain the approval from the App Store if we want to develop a mobile App. The process takes 3-4 weeks.

Considering our time and labor constraints, Geng suggested we start with a Web App, which would be easier and faster. We can ask users to bookmark and download it as a shortcut on the phone, so it can look like an App. And we don't need to worry about the phone size differences and the system design difference between IOS and Android. Geng also mentioned we should consider creating a Reporting feature in the App to collect users' feedback.

Geng's team has obtained support from the National Mentoring Partnership. One of the employees at Civic Champs has a connection with this partnership and helped them connect with the partnership. As an entrepreneur of three successful start-up projects, Geng suggested entrepreneurs establish a big network (colleagues, co-founders, friends, etc.) to obtain resources, and be ready to turn to angel investors if needed.

From the interview with Geng, we had a big picture of how to build an App from 0 to 1. We were able to set up a plan for our deliveries in the next three months. We were also able to discuss the expectations with our clients. We followed Geng's suggestions to start with a Web App. Considering there are only two members of our team who had software development experience, we decided to narrow down the scope to develop only the three major features: filter, add and edit the opportunities, which would satisfy the biggest demand of our targeted users.

Data Collection

Some of the challenges we describe above made it difficult to collect information about opportunities directly from any database. Consequently, we had to go through a manual data-gathering process. This process required visiting organizations' websites, identifying active opportunities, and scraping data about those opportunities. Our advisory board members and other stakeholders were a valuable resource in helping us identify organizations of interest.

Once we identified an organization of interest, we explored the organization's website for active opportunities that fell within 16 different categories based on HCV's suggestion. These categories include the following:

- Academic Societies
- Apprenticeships
- College tours
- Competitions
- Enrichment, Exposure, Exploration
- Honor Societies
- Internships
- Local events
- Local clubs
- Pre-college programs
- Scholarships
- Skills training
- Sports team
- Study abroad
- Summer camps
- Volunteer opportunities

Most of the categories above are self-explanatory, with few exceptions: pre-college programs and enrichment, exposure, and exploration. We define pre-college programs as educational experiences that (1) offer high school students a taste of college life by taking college courses or living on a college/university campus or (2) provide college entrance support (e.g., SAT prep). Examples of pre-college programs are Bloomfield Garfield Corporation's College and Career Readiness Program (CCRP) and Carnegie Mellon's Pre-College programs. The CCRP program supports juniors and seniors as they work to fulfill graduation requirements, explore their post-secondary education options, and navigate the financial aid process. Carnegie Mellon's Pre-College programs offer juniors and seniors the opportunity to explore their interests while taking college courses taught by Carnegie Mellon staff and faculty - mirroring the undergraduate experience.

Further, we define enrichment, exposure, and exploration programs as opportunities that provide students an introduction to a specific field, thus enriching students' understanding of their areas of interest and exposing them to pivotal resources and key stakeholders. As a result, these opportunities allow students to further explore their interests by connecting them to other opportunities. Some examples of programs that fall within this category include Carnegie Mellon University's SPARK Saturdays and Academy Pittsburgh's Beta Builder Program. Carnegie Mellon University's SPARK Saturdays introduce students to engineering concepts through hands-on activities - exposing them to engineering careers. The Beta Builder Program, on the other hand, introduces students to coding and writing software through hands-on projects.

Upon identifying an opportunity, we collected specific information about that opportunity relating to general information, logistics, and contact information. General information included features like name, brief description, whether it was a paid opportunity, and address. Logistics information incorporated cost, eligibility criteria, age range, and modality - while contact information included the name of the point of contact, email, phone number, and website.

One of the challenges we faced during this process was that some of the information we aimed to collect was not available on organizations' websites. Consequently, there were missing data for some opportunities. In order to compensate for missing data, we aimed to provide a program email or an email of the point of contact for each opportunity. We also made the decision to separate the event's location and the organization's location into different features upon realizing that these differed on occasion. In cases when they didn't, these two locations were the same in the data. Other features we added were subject - which indicated the opportunity's subject area of focus - and grade level. A data dictionary with all features is included in Appendix II.

In preparation to transfer the data stored in the spreadsheet into a database, we began data cleaning. We began to establish formatting criteria for each feature - ensuring consistency in our data. We also split selected features (e.g., address, contact name). For example, we divided addresses into separate features: address line 1, address line 2, city, state, and zip code. After data cleaning, we began the process of creating the database.

Database Creation

As we began interviewing stakeholders in the industry, we were directed toward various websites to collect details about the opportunities. We started by collecting information in an excel sheet. We intended to build the database on MySQL, and for that, we had to normalize the consolidated table of opportunities that we had initially built on Google Sheets.

Entity-Relationship Model

The first step in the data creation process was developing an Entity-Relationship (ER) model. ER modeling explains the structure of the database. It is a database blueprint illustrating how, within a system, entities (objects, people) relate with each other. We first identified entities of interest - opportunity, organization, and contact - and established their relationship.

Entity	Relationship	Entity
Organization	Hosts	Opportunity
Contact	Manages	Opportunity

We describe these relations by defining cardinality and ordinality. Cardinality states the maximum number of times one instance of an entity is associated with another. On the other hand, ordinality refers to the minimum number of times an instance within an entity is related to another. The cardinality and ordinality of a relationship comprise the multiplicity of that relationship. For example, in our data, an organization hosts at minimum one and, at most, many opportunities. An opportunity is hosted by one and only one primary organization. Adding cardinality and ordinality to our relationships above, we get the following:

Entity	Multiplicity Min..Max	Relationship	Multiplicity Min..Max	Entity
Organization	1..*	Hosts	1..1	Opportunity
Contact	1..*	Manages	0..1	Opportunity

After establishing a detailed relationship, we developed an Entity-Relation Diagram (ERD), or a visual representation of the established relationships between our entities. Once we developed an ERD (included in Appendix II), we began the normalization process.

Normalization

Normalization is a technique of organizing data - efficiently - into different tables and defining their relationships. Normalization provides multiple benefits. First, by creating different tables storing related data, we minimize redundancy - and save physical memory. Data redundancy, or duplicated data, creates maintenance issues relating to updating, adding, and deleting data. Dealing with duplicated data means that any warranted change requires someone to implement that change exactly the same everywhere. For example, any change in an organization's address is easier to implement when we store the data in a single table, an organization table than implementing it in the google spreadsheet (where organizations' names repeat across rows). Second, by reducing redundancy, every table becomes compact - minimizing memory usage. Normalization also helps improve data integrity as individuals can access and manipulate the data efficiently and quickly.

We divide the normalization process into several rules referred to as normal forms. The first normal form (1NF) is the first level of the process. Under 1NF, every table must meet three criteria

- each attribute (column) contains one type of information
- there are no duplicated attributes
- each record (row) is atomic; for example, having multiple addresses within one record violates this third criterion

Under 2NF, each table meets all criteria for 1NF and contains attributes that depend only on a single unique identifier. To meet 2NF, we created an automatically incrementing, numerical primary identifier

(key) in each table. Under 3NF, each table meets all criteria for 2NF and has no transitive functional dependencies. In other words, tables have no redundancies - where input entries once and only once.

Let's revisit the program SPARK Saturdays we discussed in the previous section. The program is SPARK Saturdays, and the subprograms are the different labs offered within SPARK Saturdays (i.e., Radio Lab, Arduino Lab). Having both program and subprogram within the same table violates the dependency criteria of 2NF. The description attribute depends on the program attribute; the program attribute depends on the sub-program attribute; the description has a functional dependency on the sub-program attribute. Consequently, to meet 3NF, we further divided the opportunities table into three tables:

- **Opportunity**; where each row is a combination of the program and sub-program with an automatically incrementing, numerical, primary key; containing opportunity details and foreign keys (primary keys in other tables - establishing a relationship)
- **Program**: where each row is a unique program - the containing program name, program email, and an automatically incrementing, numerical, primary key
- **Sub-program**: where each row is a unique sub-program -containing the sub-program name, email, and automatically incrementing numerical, primary key

While there are more normal forms beyond 3NF, creating a database achieving 3NF is common. Consequently, for this project, we stopped at 3NF. The next step in the process was creating the database on a server and adding the data to the database.

Database Creation

We utilized the ER Diagram and Lucid Chart to generate the script (Appendix II) to create the schema in MySQL. In the testing phase for the script we hosted the database on our local server, but HCV wouldn't have access to our local server post-handover. As a result, HCV purchased a server on Microsoft Azure. In order to give us access to the server, HCV created a MySQL database on the HCV server and modified the firewall permissions to allow specific IP addresses to connect to the server, create the schema, and load data into the tables. With our given set of permissions, we were not allowed to drop any tables, so we had to make sure it worked seamlessly on our server before we created the schema on the HCV server.

We utilized a number of tools to load data into the tables. We began with the Load command, but we were restricted by the firewall settings and couldn't utilize the command that would upload data directly from a CSV file. We used the Import Wizard option available on MySQL to load data in some tables, but even that failed when it came to bigger tables like the Opportunity table. For that reason, we made another attempt. We tried to load CSV data to the local database instead of the database on the server. Then we exported the insert statements based on the data we have on our local data tables. We then replaced the table name and some other attributes with Linux sed commands. After our insertion scripts were created, we copied and pasted that script and ran it on our database on the HCV server.

Web App Development

Front-end Development

We built the User Interface (UI) using React JS and configured the project using Create React App. React JS is a free and open-source Javascript library used widely to create single-page applications. To create

multiple UI components, we used Material UI. Material UI is another library that allows us to import and use different components (e.g., buttons, autocomplete, navigation bar) to build a user interface in React-based applications. We followed a Styled component structure (which lets us write CSS in Javascript) for adding styles to the components.

We followed a component's structure to create the web app that includes::

- Login screen
- Home page displaying the opportunities
- Add opportunity screen
- Edit opportunity screen
- View details page
- About page

We included the feature to filter opportunities based on grades level, regions, and opportunity type. We also added a feature for adding new opportunities and editing any existing opportunities. Whenever an opportunity is added or edited, the information is saved directly to the database with the help of REST API(described below).

Each opportunity displayed on the home page is a card view that has an option to view opportunity details. When users select the “view details” link from the home page, they are directed to a new tab where they can see the full details of the selected opportunity. We also created an about page where we added the team’s details (name, profile picture, role, and LinkedIn URL).

Back-end Development

To support add and edit opportunities features, we decided to build an event-driven server for this project. We used Node.js as the environment and Express.js as the web framework to finish server-side scripting and built several APIs to build the connection between client-side and server.

We followed the MVC design pattern to create our backend services. In the beginning, we have a script db.js that configures the connection between the server and the MySQL database. For that part, we used Sequelize as the ORM for MySQL.

```
const db = new Sequelize('hcv_data', 'admin', 'password', {
   host: "localhost",
   dialect: "mysql"
});
```

Default configuration for the database

The first parameter is the database name, second parameter is the username, and last one is the password. Further, we also have to provide the host for the server and the dialect (MySQL). Then, in the index.js file, we created an express application and built the connection between the server and MySQL using the previous settings.

After we created the connection between the database and the server, we designed the data model for the data from the table in MySQL. In the models' folder, we implemented a script called opportunitiesModel.js, which defines all the attributes' data types from a specified data table. To realize the add and edit features, we need controllers to control these logics as well. In the controller's folder, we have a file named opportunity.js. We designed four events in that file, which are getAllOpportunities, getOpportunityById, createOpportunity, and updateOpportunity.

For the getAllOpportunities function, we fetch all the data from the table and send it back to the client in JSON format. For the getOpportunityById function, we select the data based on its opportunity id. For the createOpportunity function, we received the data from the web request and used this data to insert the data into the table. And for the updateOpportunity function, we also get the data from the web request, but this time we update the table based on its opportunity id.

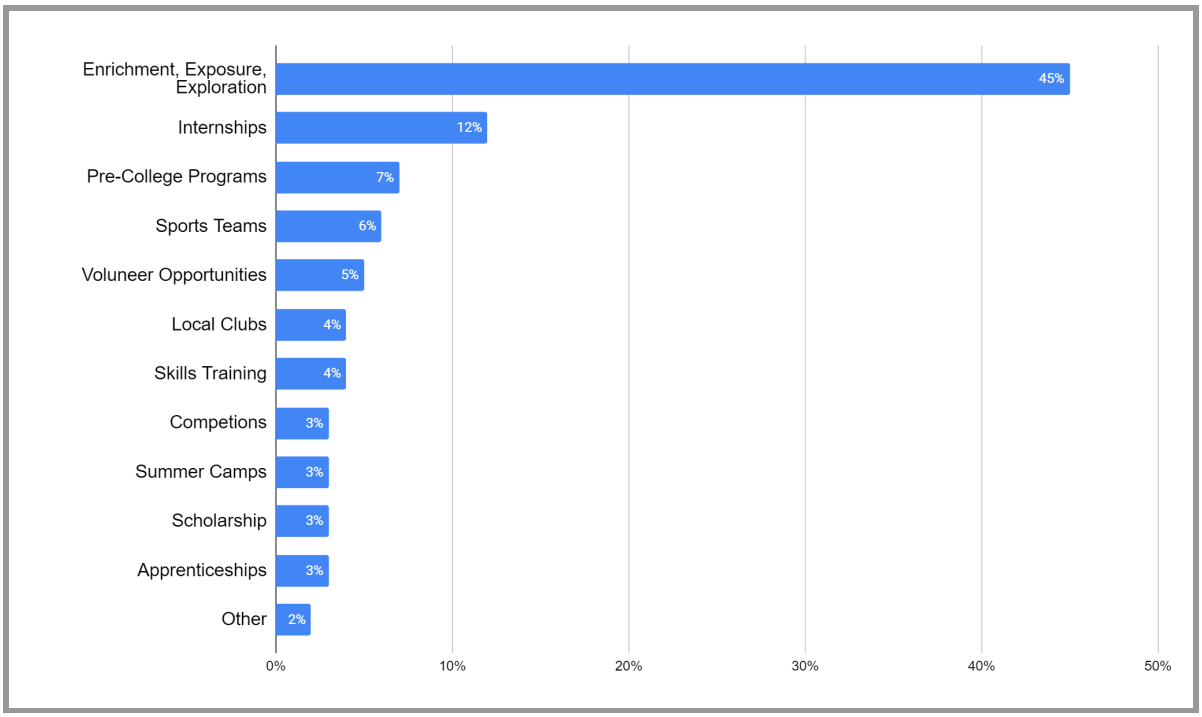
Then in the routes folder, we assigned these functions to various URL paths. In this way, the client-side can send GET, POST, or PATCH HTTP requests to the server. We tested all these interfaces using Postman to ensure the server received requests.

Data Analysis

Data Summary

By the end of the project, we collected 204 opportunities from 61 organizations. Figure 1 represents the distribution of opportunities by type. The majority of the opportunities fall within the Enrichment, Exposure, and Exploration category, followed by Internships, and Pre-College Programs. To clarify, we define enrichment, exposure, and exploration as opportunities that provide students an introduction to a specific field, thus enriching students' understanding of their areas of interest and exposing them to pivotal resources and key stakeholders. And we define pre-college programs as educational experiences that (1) offer high school students a taste of college life by allowing them to take college courses or live on a college/university campus or (2) provide college entrance support (e.g., SAT prep).

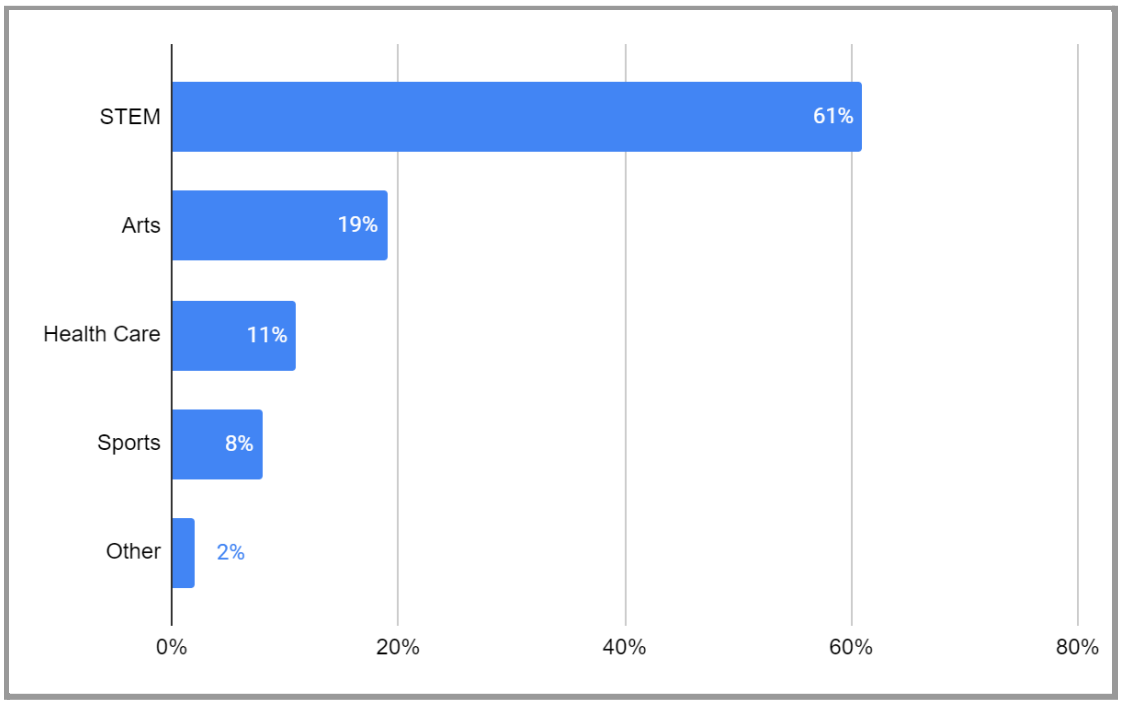
Figure 1: Distribution of Opportunities by Type



Note: Other includes Honor Societies, Academic Societies, Fellowships, Study Abroad, and College Tours

Our main focus, in the beginning, was to find STEM-related opportunities. Therefore, when exploring the distribution of opportunities by subjects, we see a heavy emphasis on STEM-related opportunities. Currently, we have a total of 24 different subjects of opportunities, which can be divided into five different categories, STEM, Arts, Healthcare, Sports, and others. The distribution of the categories is represented in Figure 2, in which 61% of the opportunities are STEM-related, followed by Arts (19%), Health Care (11%), Sports (8%), and Others (2%). For the detailed distribution of subjects, refer to Table 1 in Appendix IV.

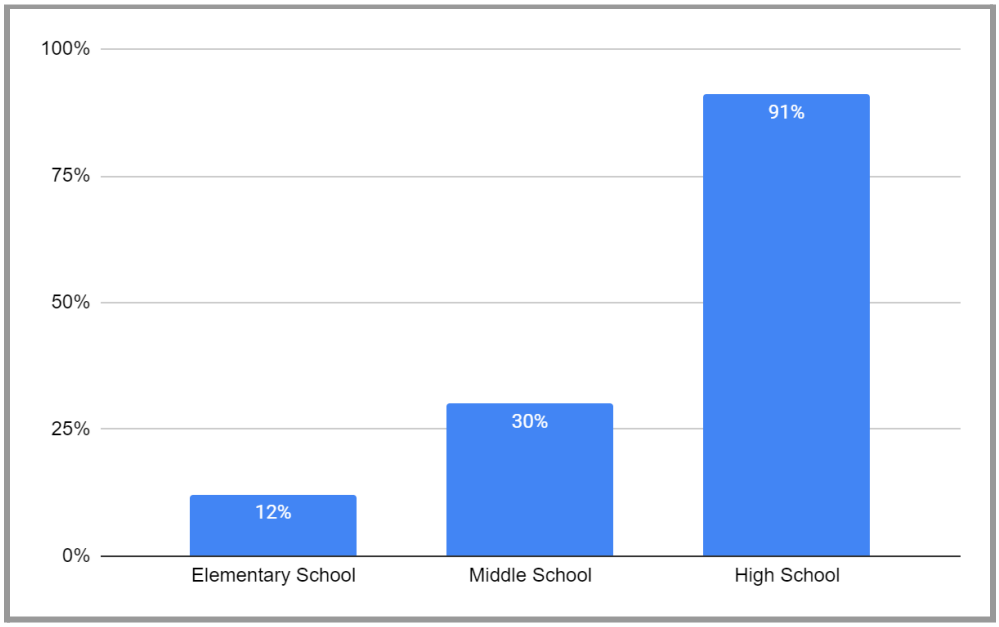
Figure 2: Distribution of Opportunities by Subject



Note: Other includes Business, Media & Communications, and Education

Our data collection focused on opportunities for high school students - which is reflected in Figure 3, with 91% of the opportunities geared toward high school students. While the target population of the majority of opportunities' are high school students, we also see that 42% of the opportunities are also available for elementary and middle school students.

Figure 3: Distribution of Opportunities by Grade Level



Note. If an opportunity is offered to students across different school levels, the opportunity is counted twice.

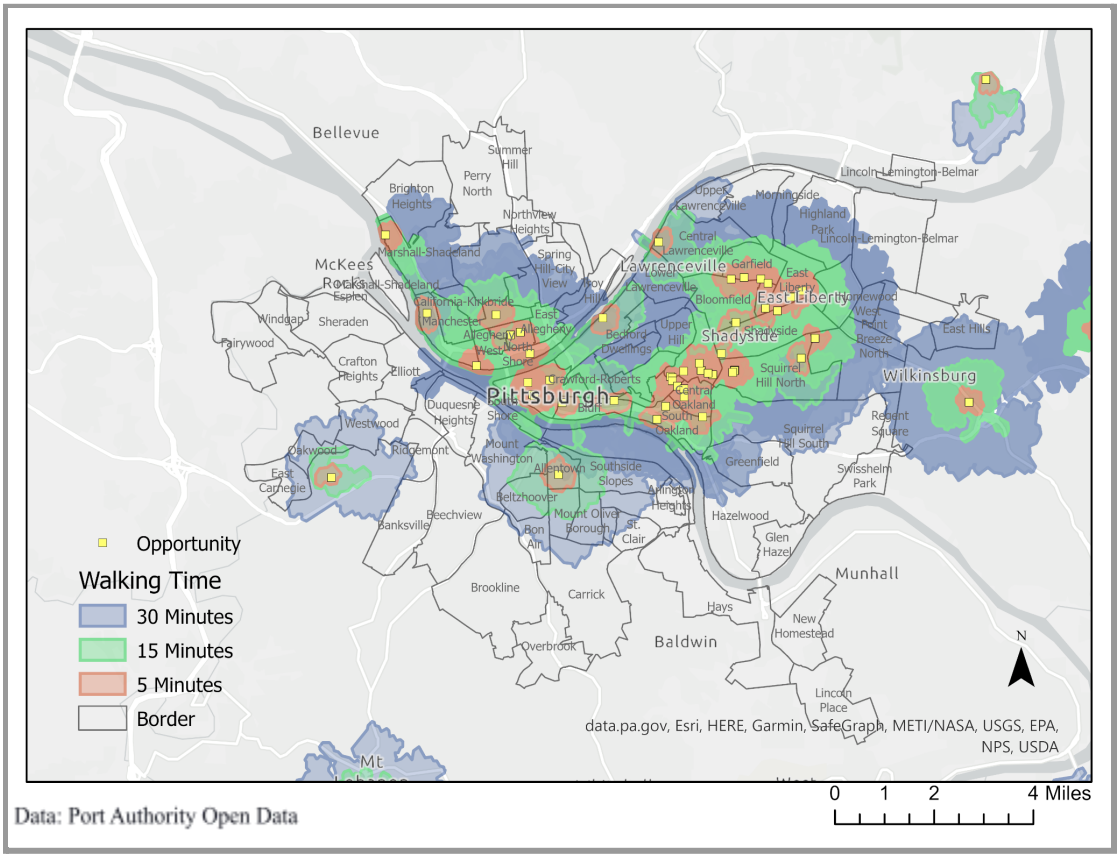
Geographical Analysis

As previously highlighted, transportation is often a barrier for children to access opportunities. While public transportation is accessible for the majority of the residents in Pittsburgh as shown in figure 7 (included in Appendix V), these options are not consistently reliable. There are such situations as heavy traffic and, especially with the bus driver shortages, there are circumstances where students need to rely on other methods such as cars or walking. Therefore, we will go over the geographical representation of the opportunities we have collected within Pittsburgh to assess its accessibility for students. ArcGIS Pro was used to create the following Maps used in the analyses.

Figure 4 illustrates the geographical distribution of the opportunities offered in person in Pittsburgh, that we have collected. The legend in Figure 4 outlines the number of opportunities in each location. The majority of the opportunities are concentrated in the Northern Pittsburgh area from Mckee Rocks to East Liberty Area. Figure 5 shows the accessibility to opportunities by walking time. In Figure 4 the opportunities seemed inaccessible for the majority of the students, but when we look at the walking time, two-thirds of Pittsburgh can access at least one opportunity within 30 minutes of walking time.

Figure 4: Distribution of Opportunities by City Level

Figure 5: Access to opportunities by walking time

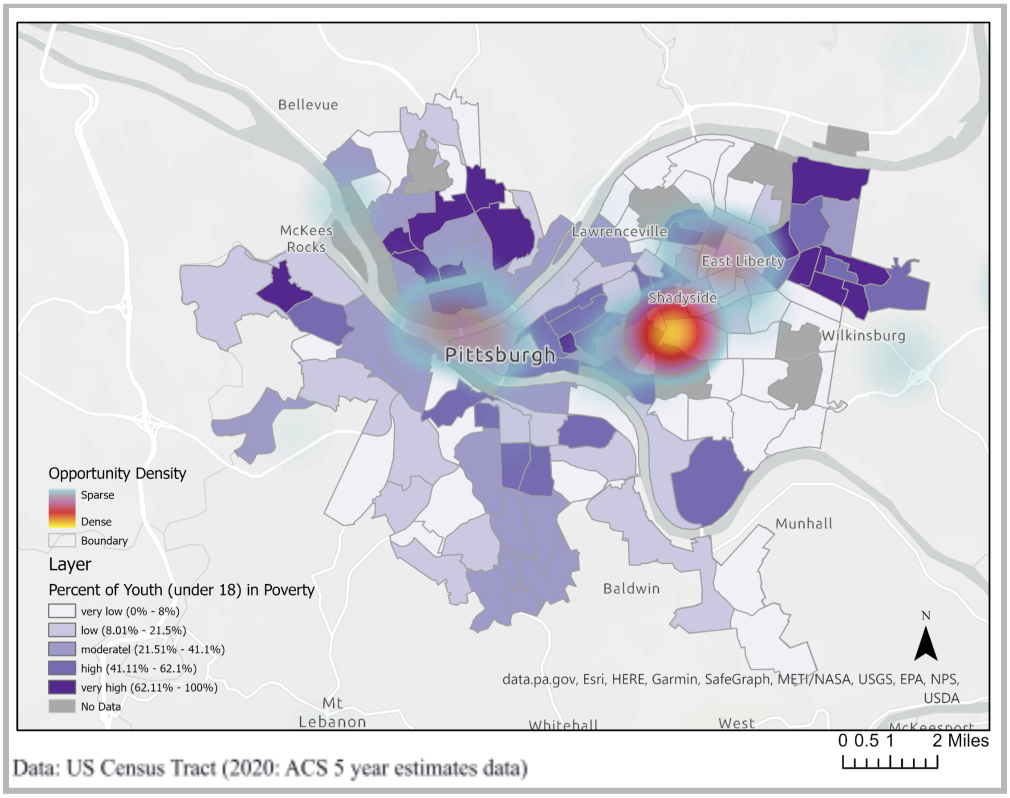


We also explored the density of opportunities and the target population, underserved youth. We defined underserved youth as youths under 18 years of age living under the federal poverty level. The legend in Figure 5 indicates the percentage of youth living in poverty and the density level of the opportunities. The percentage of youth in poverty is represented by light purple to dark purple colors. Light indicates low percentages and dark indicates high. The density level is represented by the heat map.

We initially started collecting opportunities through CMU and CMU's partner organizations. Therefore, when we look at the density level, the opportunities gathered have a heavy density in the Shadyside area where CMU, UPMC, and the University of Pittsburgh exist. Then, there is moderate density in the downtown and east liberty area.

Areas with the highest percentage of youth in poverty are centered around the Northeast (such as the highland park), Northwest (cal-bride), and western area of Pittsburgh. When these two distributions are compared, it shows the disparity between our target population and the density of the opportunities. While we do cover some of the areas with a higher rate of our target population, areas with the highest percentage of the target population are still sparse. With the limited time and resources, while we were able to expand the geographical orientation of the opportunities, areas with the highest percentage of youth in poverty are still sparse. An additional expansion of the opportunities is anticipated with the progression of the project.

Figure 6: Density of Opportunities compared to Target Population



Opportunity in the Spotlight

We would like to highlight a successful model from the 204 opportunities we have collected; High School U from the Learning Alliance. The program spans students' high school careers, starting with 10th-grade through graduation - ensuring students graduate with a guaranteed position. In high school, the program offers students paid internships, mentoring, the opportunity to take college classes, free SAT prep, and free public transportation. By providing paid internships and free bus passes the program eliminates cost and transportation barriers. The program continues to support the student through college ensuring that students continue to have professional exposure and academic, personal, and financial support.

The program began as a summer program in 2016 and in 2019 received funding to expand the program from summer to year-round. Overall, the organization managed to successfully overcome some of the organizational constraints and limitations highlighted in the Research Section.

Conclusion

Children growing up in Pittsburgh's lower-income families often lack a direct connection to the knowledge-intensive industries that are remaking the economy of our nation in the 21st century. The difficult situation is due to a lack of exposure to the opportunities, a lack of understanding of the skills required, and a shortage of mentors who can guide students into careers.

In this project, MSPPM and MISM students worked together with Homewood Children's Village (HCV) to create key components of a smartphone app that can help match interested HCV-supported students with appropriate experiences, internships, and mentorships that can help connect them to 21st-century career opportunities.

After meeting with staff from CMU's Gelfand Center, the Carnegie Museums, the Children's Museum, CCAC, UPMC, STEMisphere, and a number of other organizations, we obtained information about the needs of students and their families for CE3 experiences and services, the current status quo of the available opportunities and service providers across the region, and the challenges facing the families and the providers.

Based on the name lists obtained from various sources, we web scraped the organization data on the Internet and arrived at 204 opportunities from 61 organizations. We continued to categorize the existing opportunities into different types and subjects. To create a database, we extracted the main features of the opportunities and normalized the dataset into a ready-to-use format with the help of MySQL and Microsoft Azure. We also visualized the opportunities on maps to find the distribution patterns.

With the database at hand, we were able to create a Web App to help students and parents to connect to existing opportunities. This has come out to be a user-friendly Web App that satisfies both the demands of students and the HCV management team. Students with user accounts can easily view and filter all the opportunities on the homepage and opportunity details page. The management team from HCV can add or edit opportunities with a few clicks. The Web App also includes a Log-in page and an About page to ensure completeness.

Considering our time limits, we stopped at a point where we finished the integration of all our existing opportunities and the development of the major functionalities of the Web App. In the next step, the Web App can be tested among the pilot students and parents to evaluate its effectiveness and be further upgraded to include different roles of log-in and a mobile version.

We want to give the following recommendations for the future endeavor direction.

Recommendations

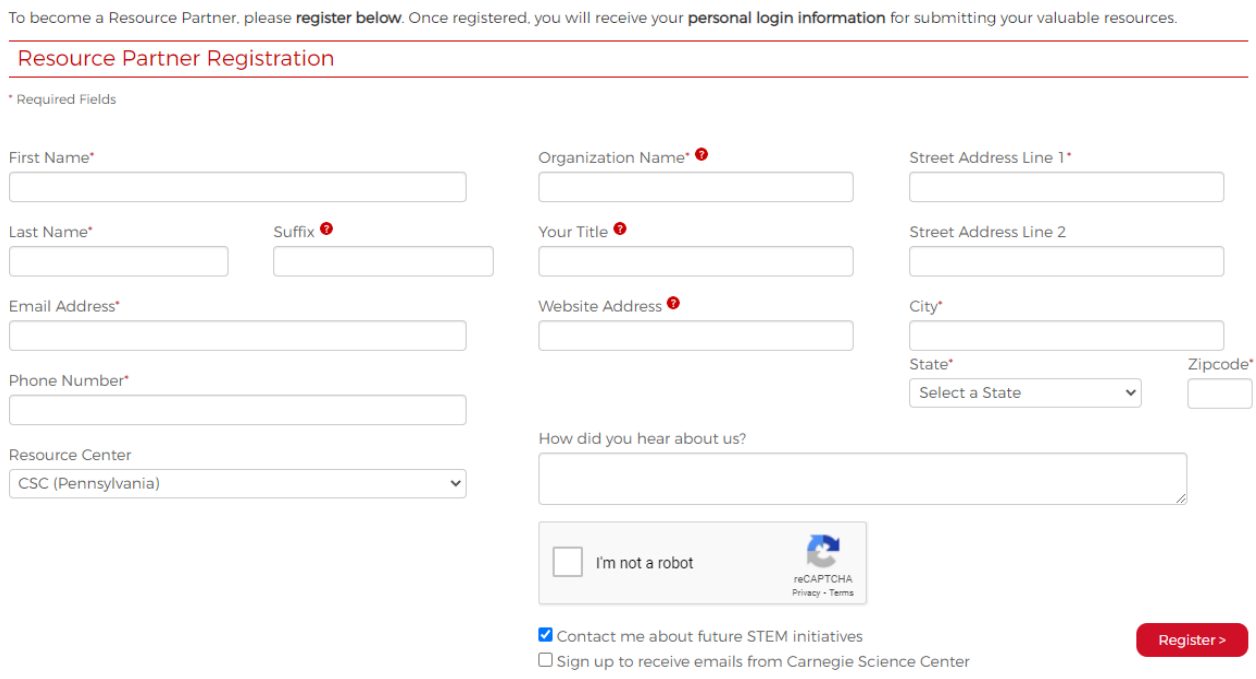
One of the advantages of working on a project in its nascent stage is that we had the opportunity of gaining a diverse experience. Through the course of our project, we talked to many people in the industry who have worked closely or observed similar projects in the field. We learned a lot from our process and their experience. While we constantly modified our project based on their advice, there were a number of recommendations that would be vital moving forward.

Shift to contributor-based model

We started our data collection process by reaching out to different organizations, conducting interviews, and collecting opportunities from them. We also requested the members of these organizations to introduce us to their counterparts in other organizations. Michele Howard from Stemisphere introduced us to their contributor-based model.

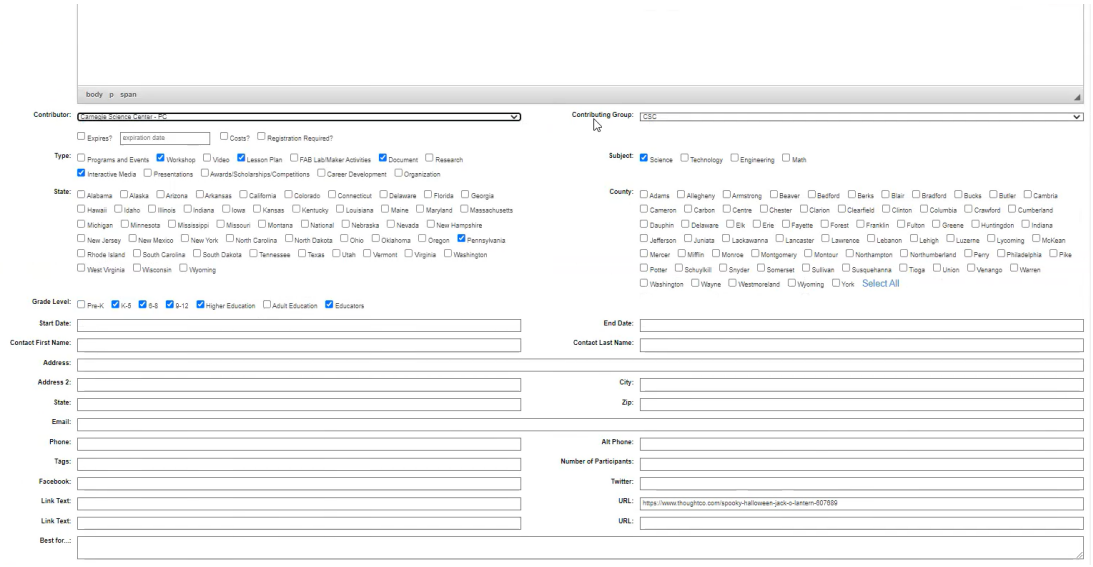
During the development stage of Future Maker's Market, data collection will be through direct conversations to the point of contact in the organization, manual web scraping, and data entry into the database. As the project evolves, we suggest HCV utilize these relationships, to encourage organizations to register as partners on the website whereby they will be able to contribute to opportunities directly. This will shift the onus of adding opportunities to the website, from the collectors to the contributors.

The website should have a tab for resource contribution where organizations can enter details and register themselves as partners with HCV, and generate a partner profile with contact information and a login and password. This will act as a logical filter for authentication. The second step would be to introduce different user profiles on the website to control access, namely Parents, Students, Partners, and HCV administrators, each user will have different access permissions. Another step that will have to be taken in parallel with the above-given steps is to prepare a rules and instructions document, which will help partners understand exactly what types of opportunities they are allowed to add to the database, and what data features will be required in order to contribute.



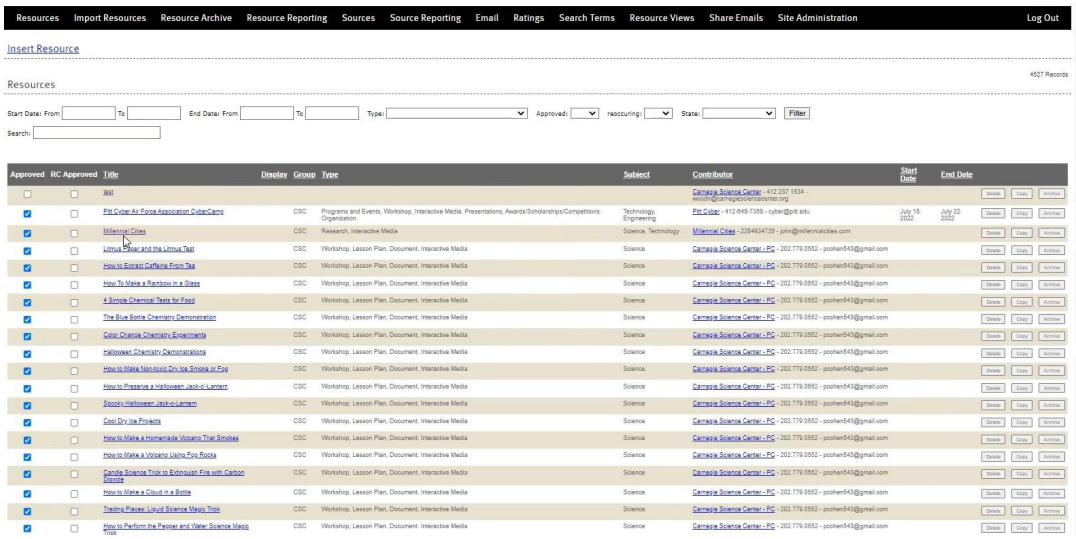
Source:"STEMisphere | Carnegie Science Center," accessed May 8, 2022, https://stemisphere.org/.

Once a partner generates a login password, they can log in to contribute. They will first be required to read the rules and instructions before they can open a data entry page. We advise making all contact-related data features as well as the URL field to be compulsory fields before they can submit - acting as a second security measure.



Source:"STEMisphere | Carnegie Science Center," accessed May 8, 2022, https://stemisphere.org/.

We also recommend developing a back end that will be accessible to the HCV database administrator. When a partner contributes an opportunity, it will get automatically added to the database, but it will not go live on the web app until the administrator authenticates and approves it.



Source:"STEMisphere | Carnegie Science Center," accessed May 8, 2022, https://stemisphere.org/.

Database Maintenance

With every interview, every conversation about this web application, with anyone in the industry, we heard about one obstacle over and over again. Similar endeavors have lost momentum because organizations are unable to maintain an up-to-date database.

Database maintenance is a resource-intensive process. And while shifting to the contributor model will reduce the burden, there are many other aspects of the database that will require regular maintenance. We advise that a yearly authentication cycle should be established to ensure that all the opportunities live on the website are still active and the requirements and eligibility are still valid. When we started our outreach to these organizations, we documented the process that HCV can customize to its model. Having an established outreach process will enable volunteers and interns to contribute to Future Maker's Market.

On the backend, there will also be a need for technical support. For instance, when a contributor adds an opportunity, the information will go to different tables, each of which has a primary key that is also reflected as a foreign key in the Opportunity table. While we can program the primary key to be auto-incremented, the foreign key value for each entry in the Opportunity table has to be done manually. This is just one instance of the technical support that Future Maker's Market will require. HCV can hire dedicated resources for the purpose or can hire a third party for technical support.

Design a Success Matrix

Another critically beneficial feature would be to define success metrics, which will not only help Future Maker's market become a prominently used tool in the community but also help HCV decide where to focus their resources. Here are some metrics that can be used as parameters for measuring success. HCV can use tools like google analytics or Hub spot to analyze all the information and draw useful conclusions.

The number of visitors: It measures the number of people that visit the website in a given month. It will be an indicator of the popularity and effectiveness of the website. In case of low traffic, HCV can modify the content to rank higher in search engines or run ad campaigns on Google Ads, and other social media platforms. If there is well-established traffic, then HCV can measure the return visitor metric, which will show the number of visitors that returned to the website at least once. This will be an indicator of how engaging the content is.

Interactions per visit**:** It shows how users move across the website, which parts of the website are most engaging, and which aren't working as efficiently as they were meant to. It also allows you to see the exact link, button, or other interactive elements a visitor has clicked on. This parameter will tell you which opportunities are popular amongst the students and have high traffic. HCV can focus its efforts on collecting more of that particular category of opportunity.

Bounce Rates: A website's bounce rate measures how many visitors leave a page without performing any specific action, or engaging with the website. Having a high bounce rate can make your website appear lower in search results. Knowing which pages of the website have the highest bounce rates will tell you which opportunities are not so popular. HCV can then focus its resources on opportunities that have higher traffic.

Device Sources: The device sources will point out what kind of devices users accessed the website with. It can help HCV better target the expansion plan. If they were to turn Future Maker's Market into an app,

device sources will inform the decision on which operating system should they target first - IoS or Android.

Feedback and Rankings: The web app also features the option for users to rank or give feedback for any given opportunity. This information should also be added to the database and will require a separate table which will have a composite primary key of the opportunity id and the date of feedback. This feature will allow the user to review the opportunities that they have accessed via Future Maker's Market and also provide rankings to signify the quality of their experience with that organization.

Awareness Program

From our conversations with industry experts we understood that reaching the target audience is a major obstacle with respect to a product like Future Maker's Market. Keeping that in mind, we have created a preliminary design of an awareness program.

In our conversation with Jackie Foor, the executive director at Consortium for Public Education, we learned that students are reluctant to follow advice from the adults in their lives. It is instead more effective to have students be peer mentors. We suggest involving students who have used the Future Maker's Market and accessed fruitful opportunities, in spreading awareness in the community. Bring them in to tell their success stories, and make them the *brand ambassadors*. Another thing we suggest is to host fundraiser events. Fundraisers will serve the dual purpose of raising funds for the expansion plan and providing a forum to publicize the contributor-based models. These fundraisers can be used as an opportunity to establish a network of partner organizations.

We also learned that Pittsburgh has reached an *inflection point*. There has been an unprecedented level of baby boomer retirements while there is a lack of a deep pipeline of younger talent. Over the next ten years, 1.2 million workers will need to be hired or upskilled in Pittsburgh. This is reflected in the increasing interest by organizations in training and mentoring students. HCV can take advantage of the inflection point, by becoming a conduit for these organizations in preparing the future workforce of Pittsburgh. We would also like to suggest building *partnerships with schools*. There were a number of organizations that had excellent mentoring programs, but they only partnered with schools. If HCV gains access to these schools, they can interact with other stakeholders as well as increase the user base of the website.

Logistics and Safety

In our conversation with the Consortium for Public Education, we learned that one of the biggest barriers in connecting mentors with schools was ensuring child safety. While it helps that schools have access to the federal database, and do thorough vetting before allowing contact, that is not the case with Future Maker's market. We advise HCV to design a vetting and authentication process that ensures child safety when students access these opportunities. The partner registration feature of the website will be the first step in this direction.

Another important lesson we learned was that while the website will give access to opportunities, there still remains a resource gap, especially in the case of underserved children, for instance, lack of access to transportation. There are a number of opportunities that lie beyond the reach of public transport. And most of our opportunities are in STEM-related fields, and while covid has transformed education in Pittsburgh by providing access to laptops and the internet, there still remains a sizable gap.

Login & Authentication
Currently we have not included Authentication to ensure the type and identity of users logging in. Adding Login & Authentication such as 2-FA will provide added security by allowing role-based login. It will also ensure the correct user is logging in and hence, it will provide a secure platform for its users.

Add features for different roles
We recommend adding different roles such as Students, Staff, Mentors, Partners, and Family to the log-in screen. For example, students can have access to all available opportunities but they will not have permission to add or edit new opportunities. Opportunities can be added or edited only by a specific role such as HCV staff. Likewise, additional features can be added based on different roles.

Roll out an app to Pilot users
After an initial level of testing, HCV can roll out the app to pilot users that may include a certain set of users i.e students/partners/mentors. This will ensure the app is functioning well without any issues before the actual launch.

Mobile app development
Today, mobile apps are essential for expanding the business reach. The current web app can be enhanced in the future to support mobile devices to gain massive exposure and reach a wider audience. This will help HCV stay ahead of the curve.

References

1. “9 Important Website Metrics You Should Track,” HostPapa Blog (blog), November 5, 2019, https://www.hostpapa.com/blog/analytics/9-important-website-metrics-you-should-track/.
2. “STEMisphere | Carnegie Science Center,” accessed May 8, 2022, https://stemisphere.org/.
3. “Allegheny Conference —,” accessed May 8, 2022, https://www.alleghenyconference.org/beyondinflectionpoint/.

Appendices

Appendix I - Outreach Process

Upon conversing with Homewood Children's Village and understanding the organization's goals concerning this project (gather opportunities for their students and secure future collaborations with other organizations), the Systems team created an outreach process to ensure the ultimate goal is reached. This outreach process was executed as we held conversations with advisory board members and stakeholders, and it is outlined below –

Pre-Interview Process

Before each interview, a designated member of the team prepares a *one-pager* that contains background information of the person interviewed, their position, their organization, and any relevant information. This one-pager is of primary use to the systems team, as it serves as a point of reference for members to prepare for the interview. The team also prepared a *questionnaire document* that includes a list of questions to be asked during the interview.

Interview Structure

During the interview, the team assigns a spokesperson who asks the questions prepared and a note-taker who fills out the questionnaire document with the answers provided. Other members ask follow-up questions when necessary. After the guest has finished answering questions, the spokesperson inquires about other resources the guest can provide. The team always asks for at least *three points of contact* that we can follow-up with to continue the conversation, and we ask if the guest is willing to make the introductory connection via email. Additionally, the team inquiries about *potential follow-up conversations* with the guest, especially since we noticed that one hour tends to be a short amount of time to obtain all the information we are looking to learn for the guest and their organization. Furthermore, the team inquiries about *establishing a connection between the guest's organization and HCV*. As HCV made it clear that part of the goal is to build connections with other organizations, the team made it a priority to inform guests about HCV's desire to collaborate with, and offer to bridge the gap by introducing the guest to Anton via email. Lastly, the team inquires about having a HCV representative in any follow-up meetings that we may hold with the guest and/or any member of their organization. This step was simply to build on the previous step and facilitate any future collaborative efforts between both organizations.

Post-Interview Process

After meeting with our guests, a designated team member sends a follow-up email to the guest to thank them for their time, to provide the team's availability for any follow-up meetings (if necessary), and to obtain points of contact of any resource provided to us by the guest during our conversation. The team also sent Lee a summary of the lessons learned during our conversation, which was later discussed during weekly meetings with Lee and Anton. Any new information learned during our interviews was constructed as a question for subsequent meetings to determine if the information was persistent among organizations.

Appendix II - Database Documents
Data Dictionary

Organization Table

This table contains one row per organization, containing the organizations' names, addresses, and unique organization ID that is incremented by one.

Feature	Data Type	Data Description
org_name	varchar	Organization Name
org_address_line_1	varchar	Organization Street Address
org_address_line_2	varchar	Organization Building Name/Suite Number/...
organization_city	varchar	Organization City
org_state	varchar	Organization State in Abbreviation
org_zip	integer	Organization 5 Digit Zip Code
org_id	integer	Manually created ID

Program Table

This table contains one row per program, containing the programs' names, email addresses, and unique program ID incremented by one. This table contains the main programs. For example, CMU's pre-college program is the main program. The different programs (Drama, Music) are subprograms- noting that not all programs will have subprograms.

Feature	Data Type	Data Description
program_name	varchar	Name of the program
program_id	integer	Manually created ID
program_email	varchar	Program Contact Email
organization_id	integer	Foreign Key, establishing a relationship with Organization Table

Sub-program Table

This table contains one row per subprogram, containing the programs' names, email addresses, and unique subprogram ID incremented by one.

Feature	Data Type	Data Description
subprogram_name	varchar	Name of the opportunity
subprogram_id	integer	Manually created ID

program_id	integer	Foreign Key establishing a relationship with the Program table
subprogram_email	varchar	Email for subprogram

Event Location Table

This table contains one row per event location, per opportunity, and per date - containing the event address and unique event ID incremented by one. The date feature in this table will reflect any changes in event locations for a single opportunity. More specifically, if the event address of an organization changes, then a new row is added to the table. The date in this example will reflect when this change occurred.

Feature	Data Type	Data Description
event_address_id	integer	Manually created ID
event_name	varchar	Event of name
event_address_line_1	varchar	Event Street Address
event_address_line_2	varchar	Event Building Name/Suite Number/...
event_city	varchar	Event City
event_state	varchar	Event State in Abbreviation
event_zip	integer	Event 5 Digit Zip Code
opportunity_name	varchar	A combination of program:subprogram
date	date	the date information was added

Contact Table

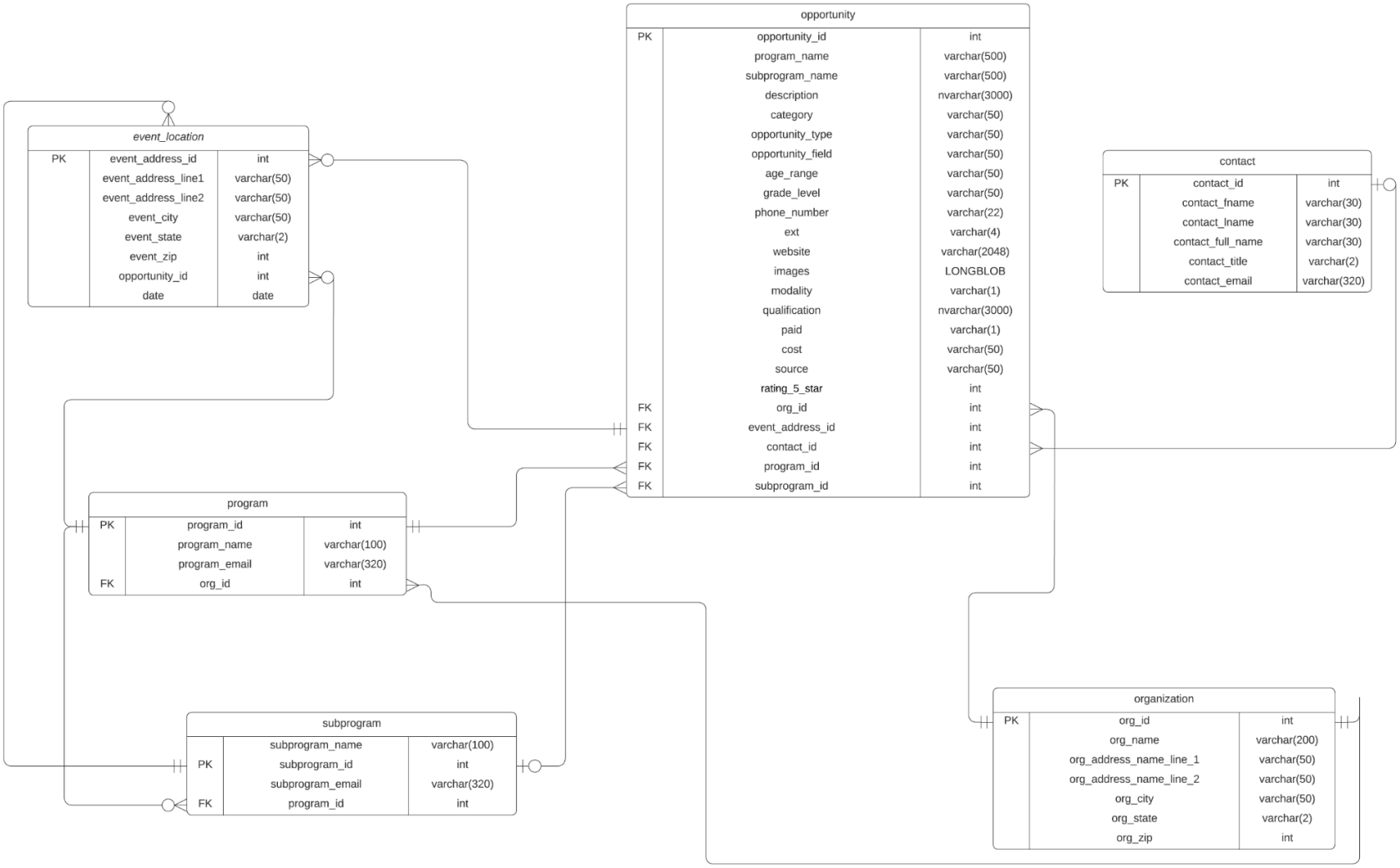
This table contains one row per point of contact - containing the contact name, contact title, contact email, and unique contact ID incremented by one.

Feature	Data Type	Data Description
contact_id	integer	Manually created ID
contact_title	varchar	Contact Person's Title
contact_fname	varchar	Contact Person's First Name
contact_lname	varchar	Contact Person's Last Name
contact_full_name	varchar	Contact Person's Full Name
contact_email	varchar	Contact Person's Email

Opportunity Table

Feature	Data Type	Data Description
opportunity_id	integer	Manually created ID
program_id_	integer	Foreign Key
subprogam_id	integer	Foreign Key
contact_id	integer	Foreign Key
organization_id	integer	Foreign Key
event_address_id	integer	Foreign Key
description	varchar	Description of Opportunity
category	varchar	Type of opportunity
opportunity_field	varchar	Subject of focus
opportunity_type	varchar	Type of opportunity - granular level
age_range	varchar	Age range the opportunity targets
grade_level	varchar	Grade level the opportunity targets {Elementary, Middle, and High School}
phone_number	varchar	Opportunity Phone Number
Ext	varchar	1-4 digits internal number used within a PBX system.
website	varchar	Website URL
Images	longblob	Images for the Opportunity
modality	binary	Indicators of opportunity modality {0 - in person; 1 - virtual; 2 - hybrid; Null-no information}
qualification	varchar	Opportunity Qualifications/Requirements
paid	binary	Binary indicator of whether the opportunity is paid
cost	varchar	How much it would cost to attend the opportunity
source	varchar	Information Source
reviews	varchar	Comments from previous participants
rating_5_star	integer	Average rating score from previous participants

Entity-Relationship Diagram



Script

Database Creation Script:

```
CREATE TABLE `organization` (
  `org_id` int,
  `org_name` varchar(150),
  `org_address_name_line_1` varchar(50),
  `org_address_name_line_2` varchar(50),
  `org_city` varchar(50),
  `org_state` varchar(50),
  `org_zip` int,
  PRIMARY KEY (`org_id`)
);


CREATE TABLE `program` (
  `program_id` int,
  `program_name` varchar(500),
  `program_email` varchar(320),
  `org_id` int,
  PRIMARY KEY (`program_id`),
  FOREIGN KEY (`org_id`) REFERENCES `organization`(`org_id`)
);


CREATE TABLE `subprogram` (
  `subprogram_name` varchar(500),
  `subprogram_id` int,
  `subprogram_email` varchar(320),
  `program_id` int NOT NULL,
  PRIMARY KEY (`subprogram_id`),
  FOREIGN KEY (`program_id`) REFERENCES `program`(`program_id`)
);

CREATE TABLE `contact` (
  `contact_id` int,
  `contact_fname` varchar(30),
  `contact_lname` varchar(30),
  `contact_full_name` varchar(30),
  `contact_title` varchar(30),
  `contact_email` varchar(320),
  PRIMARY KEY (`contact_id`)
);
```

```
CREATE TABLE `event_location` (
  `event_address_id` int,
  `event_address_line1` varchar(50),
  `event_address_line2` varchar(50),
  `event_city` varchar(50),
  `event_state` varchar(2),
  `event_zip` int,
  `opportunity_name` varchar (500),
  `date` date,
  PRIMARY KEY (`event_address_id`)
);


CREATE TABLE `opportunity` (
  `opportunity_id` int,
  `description` nvarchar(3000) NOT NULL,
  `category` varchar(50),
  `opportunity_type` varchar(50),
  `opportunity_field` varchar(50),
  `age_range` varchar(50),
  `grade_level` varchar(50),
  `phone_number` varchar(22),
  `ext` varchar(4),
  `website` varchar(2048) NOT NULL,
  `images` LONGBLOB,
  `modality` varchar(1),
  `qualification` nvarchar(3000),
  `paid` varchar(1),
  `cost` varchar(50),
  `source` varchar(50),
  `rating_5_star` int,
  `org_id` int NOT NULL,
  `event_address_id` int,
  `contact_id` int,
  `program_id` int NOT NULL,
  `subprogram_id` int,
  PRIMARY KEY (`opportunity_id`),
  FOREIGN KEY (`subprogram_id`) REFERENCES `subprogram`(`subprogram_id`),
  FOREIGN KEY (`org_id`) REFERENCES `organization`(`org_id`),
  FOREIGN KEY (`contact_id`) REFERENCES `contact`(`contact_id`),
  FOREIGN KEY (`program_id`) REFERENCES `program`(`program_id`),
  FOREIGN KEY (`event_address_id`) REFERENCES `event_location`(`event_address_id`)
);
```

Appendix III - Web App

This is the link to the github repository that contains web-app codes and additional resources:
https://github.com/Preity24/hcv-app.git

Appendix IV - Distribution of Opportunities

Table 1: Detailed Distribution of Opportunities by Subject

Subject	Count
STEM	57
Arts	35
Engineering	21
Health Care	20
Sports	17
Architecture	13
Technology	12
Computer Science	11
Environmental Science	7
Robotics	4
Cybersecurity	4
Biology	3
Business	3
Biomedicine	2
Linguistics	1
Nutrition and Exercise	1
Artificial Intelligence	1
Drama	1
Music	1
English	1
Education	1
Creative Writing	1
History	1
Media & Communication	1

Appendix V - Map

Figure 7: Public Transit Accessibility in Pittsburgh