Spring 2025: Rice D2K Capstone Project Proposal

Email your Project Proposal to d2k@rice.edu for review **before May 3, 2024**.
Contact us ahead of time if you wish to engage in discussion regarding your proposal.

- Review our Rice D2K Capstone **Fall 2024 Brochure and Timeline →**
- Check **Rice D2K Capstone Fees →** to support the Capstone Program and become a D2K Member to avail various partnership benefits for your organization.

Project Title

Please provide appropriate project title

Making education data and research accessible to Texas school districts – Version 2.0

Project Pitch

Please provide 2-3 paragraphs on the background and significance of your project. Include the impact your project will have on your organization, along with 1-2 visuals to illustrate the opportunity. This project pitch should be appealing to students.

In an era of increasing pressure on public school districts to be meeting accountability standards, Texas school districts are often trying to use data to drive their decision making. However, school districts often have small, or non-existent, research offices, meaning they do not have the time or capacity (or at times, analytic capability) to download, clean, and analyze these data which are made available for their benefit.

The Houston Education Research Consortium (HERC) at Rice University, works with 8 Houston area districts to ensure they have free access to high quality research. But, some of the types of questions they have are simple, straightforward things they could answer themselves, if they had the ability to access and analyze data through an intuitive web-based tool.

The D2K team who worked with HERC in the fall of 2024 did a great job of collecting data and starting to develop a tool for school districts to use. There are two primary areas a new team could build on this work: 0) Clarifying and automating the webspider/ data retrieval activity so that we establish a re-useable tool set to maintain the resource with new data 1) in identifying other analytic strategies for identifying "similar" school districts, 2) in further refining an interactive tool so that a district could also decide for themselves what characteristics are important to them in identifying a similar school.

Project Description

Please provide 1-2 paragraphs describing your project and the scope of work you envision.

As the students in Fall 2024 created a first pass at a platform housing data for districts to download and/or analyze, we believe there is an opportunity to further refine the tool by

giving districts greater ability to identify districts who are similar to themselves. Importantly, this would include the being able to identify/select characteristics that they think are important in identifying similar districts. For example, different districts may care differently about which demographic characteristics (racial/ethnic composition, language composition, economic disadvantage composition, size, etc.) or other behavioral characteristics (drop out rates, graduation rates, test scores, etc.) are important to them in identifying a "similar district". When districts can see a district similar to themselves, which has different outcomes for students – such as a higher test scores – it can be an opportunity for the districts to know who they may want to reach out to, to learn from.

Last semester's team took a pass at using an analytic strategy to match similar districts . A question remaining is: are there other analytic strategies that could be used for this exercise? Also, can there be an opportunity for districts to decide which characteristics are important to them in determining what a "similar" district is (maybe someone could check the boxes of the characteristics they want included in creating a list of their "matched"/similar districts)

Project Objectives

Please provide bullet points outlining your project objectives.

- Objective: To create a data assembly toolkit to extract and update data from enumerated public sources. This is partially done but should be evaluated as an objective.

- Objective: To create an online data tool Texas public school districts could use for accessing data and conducting simple analyses
- Objective: To speak with Houston-area school districts around what their needs might be for a data analytic tool using publicly available state education data
- Objective: To help school districts identify similar districts, as a means of finding peer districts who they might be able to connect with/learn from
- Objective: To conduct some initial analyses demonstrating the capabilities of the tool

Project Duration

Do you consider this as a one-semester project or a two-semester project?

One-semester project; as a follow up to the fall semester

Project Outcomes

Please provide your plan with respect to project outcomes and artifacts (such as codes, models, reports, etc.). Do you expect to own or protect the IP, file patents, commercialize, continue to internally work on the project in your organization, publish it as a part of peer-review article, submit it as a part of grant proposal, etc.?

We hope to make any code we create publicly available. HERC sees the work we do with school districts as a public good, and as such, want to make sure what we create is available to many audiences at no cost to them.

In terms of other outcomes, it would be helpful to have some type of web-based tool that school districts would be able to access to synthesize the data which is collected, conduct some analysis of their own, and identify other districts who are similar to them.

Finally, in being able to create a way to categorize what it means to be a "similar" district, this measure may be something which HERC can use in future projects in which it would benefit districts to understand how they are doing in comparison to other, similar districts from around the state. Additionally, this similarity measure may help districts find peer districts, who seem similar to them (demographically) who may be having different outcomes for students (higher attendance rates, fewer behavioral issues, increased academic performance, higher graduation rates, higher post-secondary preparation, etc.), who may be a source of learning.

Data Description

In the space below, describe the data for your project in 1-2 paragraphs. Please describe how the data is arranged and structured as well as the approximate size of the data. Please provide complete list of the data fields/features present in your data

The data for this project are all publicly available on TEA's website. The data are often in files which include one school district, or one campus, and one year of data, or one data topic (accountability scores OR attendance OR graduation, etc.). The D2K team who worked with us in the Fall 2024 already downloaded many of these files. It would be helpful for this Spring's team to see if there is a different way to organize the data, or whether they have any ideas about ways to automate the process in case we want to update in future years.

Data Confidentiality

Answer the questions below about the data you will be providing for the project.

1. Is your data confidential? Does it have privacy and/or security requirements to be considered? Does data have Protected health information (PHI) that is necessary for the project?

No, data are not confidential or have PHI.

2. By what mechanism will you be sharing the data with Rice? (e.g. shared via BOX folder/Google Drive/Microsoft Teams on Rice server, or kept on a sponsor-controlled server, or held on a preferred secure third-party platform). How will the students access the data?

Students will access data directly from the Texas Education Agency (TEA) website. This is the state department of education.

3. This is a course at Rice University and students will need to present their work.

   a. Public presentations. Student teams will present their findings at the end of

semester D2K Showcase in the form of a 1 minute elevator pitch video as well as a short presentation and poster session. No Confidential Information will be included in the presentations. The 1 minute elevator pitch video will also be posted publicly on the D2K website.

Do you consent to public presentation of this project?

X   Yes

❑   No

b. Internal presentations. In class, student teams will present and discuss their findings with instructors and with other students. No Confidential Information will be included in

the presentations.

Do you consent to the internal presentation and discuss of this project in a class setting?

X   Yes

❑   No

c. **Do you consent to public disclosure (title and one sentence description) of this project on the D2K website and social media?**

X   Yes

❑   No

Student Teams

Describe any special requirements or skill sets that are required from students.

Project Type: e.g. Healthcare/Social Impact/Industry, etc Social science/public education	Student Restrictions: Leave Blank
Technical Area(s): List...	Student Preferences: List major...

Skill Sets

- **Required:** Web scraping (?); Data cleaning; Data analysis (basic descriptive methods, correlational methods possibly)
- **Preferred:** Interest in developing communication skills with external partners (hearing from school districts, translating that into final product)

Sponsoring Organization

Tell us about yourself and/or your organization and its business in a paragraph below. Is your organization For-Profit or Non-Profit?

The Houston Education Research Consortium (HERC) is a research-practice partnership, which is part of the Kinder Institute at Rice University. We have conducted research, for free, in partnership with Houston area school districts for 13 years. All of the research we do is driven by the needs of school districts, around decisions they are trying to make, so that the research is ultimately used.

Sponsor Mentors

Sponsor mentors meet with the student team regularly (e.g. weekly) during the course of the capstone to guide, answer queries, and offer advice on progress and deliverables. Please list the full name, contact information, and a brief bio for each sponsor mentor.

Dr. Erin Baumgartner

Emb10@rice.edu

371 Kraft Hall

https://profiles.rice.edu/faculty/erin-baumgartner

Sponsor Administrative, Communications & Financial Contacts

Please list the names and contact information for administrative, communication/public relations and financial contacts. We want to work with your communications team to promote your organization and our partnership through social media and approved news articles. For organizations signing a sponsored research agreement include the name of the person we may contact.

Rice Faculty Mentors

Include any history of collaboration with Rice faculty.

Dr. Baumgartner has partnered with the SOPA program, working with students (and affiliated faculty) on capstone projects for 3 years. She is an instructor in the Master of Social Policy Evaluation program at Rice. Dr. Ruth Lopez Turley, director of the Kinder Institute, is a professor of Sociology at Rice.

Data Science Education

Following the completion of the capstone project, would you be interested in D2K faculty instruction in data science and AI aligned with your organization and capstone project?

- ❏ ***Extremely interested*** - *Highly enthusiastic about enhancing our team's skills in data science and*
  AI and are eager to start as soon as possible.
- ❏ ***Very Interested*** - *See great value in this opportunity and are strongly considering it.*
- ❏ ***Slightly Interested*** – *Have some interest but need more information to decide.*
- ❏ ***Not interested*** - *Do not see this as a fit for our team's needs at this time.*

Describe your interest and potential impact of above on-site capstone-integrated AI education offering on your team and business.

Is your project funded?

If yes, please provide us details and any weblinks to the grant/award/article.

No