Using an AI Agent to Facilitate Student Data Literacy and Exploratory Data-Based Investigations for All

Daniel Damelin, Senior Scientist, The Concord Consortium
Geoff Freed, Director, onfocusAccess Consulting

November 21, 2025

The Concord Consortium

Introduction and Project Goal

Project partners

- The Concord Consortium
- onfocusAccess Consulting
- Special consultants
  - Sina Bahram
  - Vikram Kumaran

This material is based upon work supported by the National Science Foundation under Grant No. DRL-2418281. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

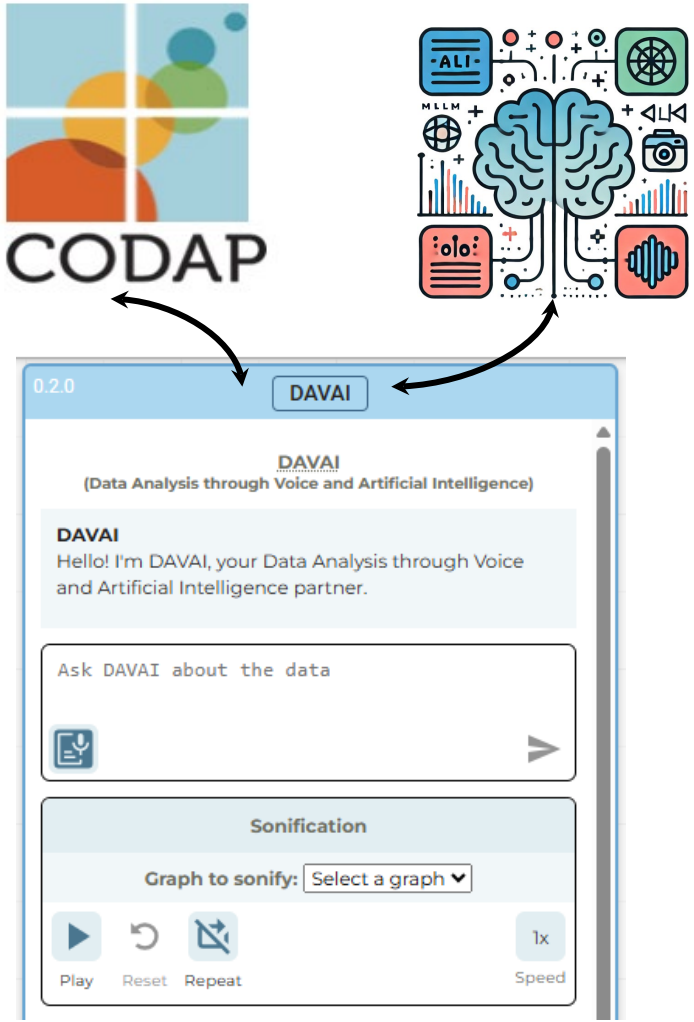
Overarching Goal:

Facilitate data exploration and the learning of data science concepts for students who are blind or have low vision.

The Concord Consortium

Reaching Our Goal

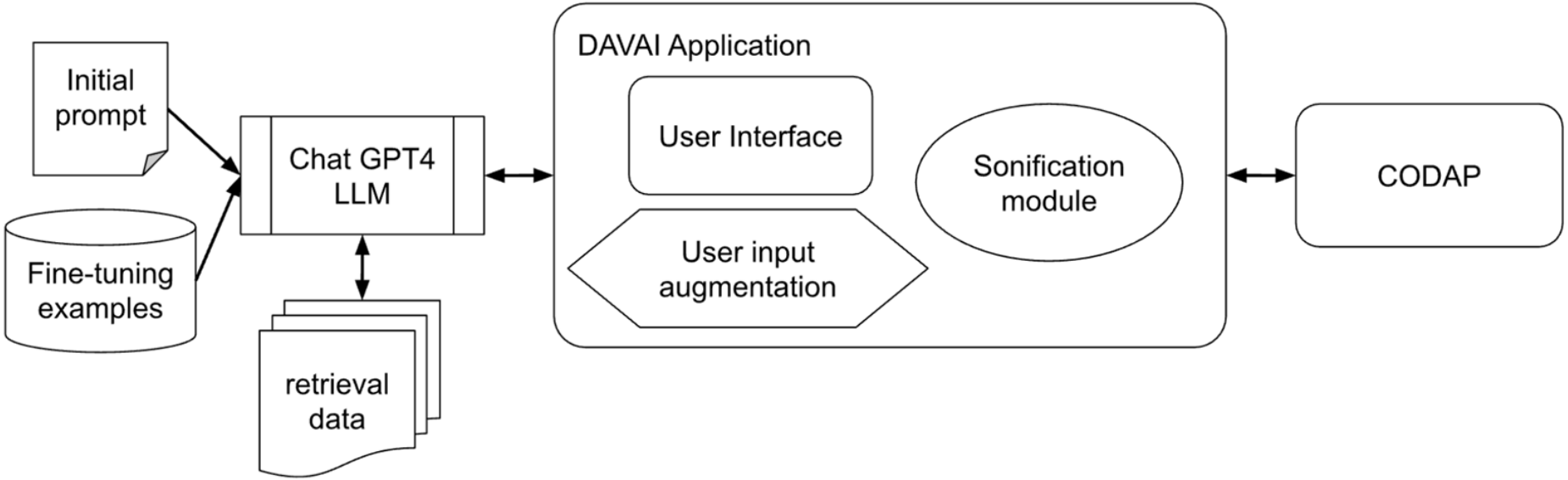
- Build on previous work with the Common Online Data Analysis Platform (CODAP)
- Create a plugin that forms a bridge between a multimodal large language model (MLLM) and CODAP
- Iterate multiple times based on student, staff, and consultant feedback
- Research technical limitations and student learning



CODAP Background and Demonstration

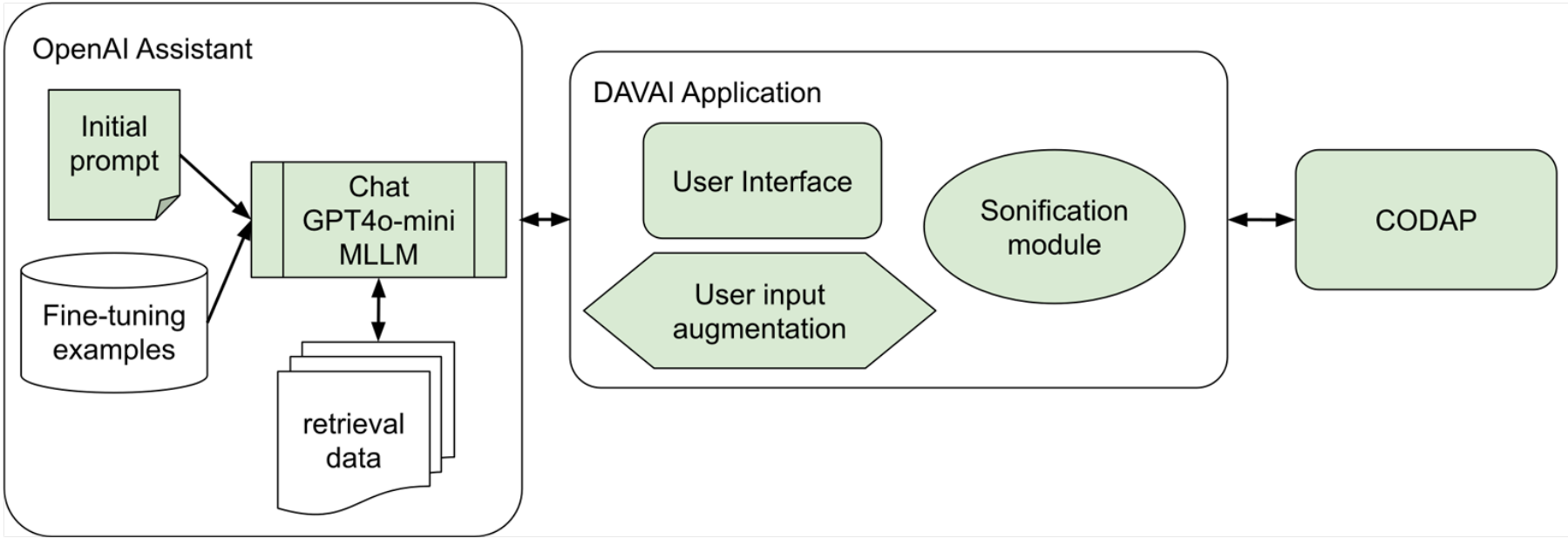
- Based on statistics education R&D going back to early 1990s
- Predecessors: desktop-based data exploration environments Fathom and TinkerPlots
- CODAP emerged through co-design-based research in collaboration with projects at Concord, EDC, and the Minnesota Population Center starting in 2014
- An immersive, exploratory data environment designed for students in grades 5–14, runs entirely in a browser, free, classroom-friendly
- Features drag-and-drop creation of maps and graphs, dynamically linked selection, and hierarchical data structures
- Extensible through plugins — DAVAI is an example

Initial Architectural Plan

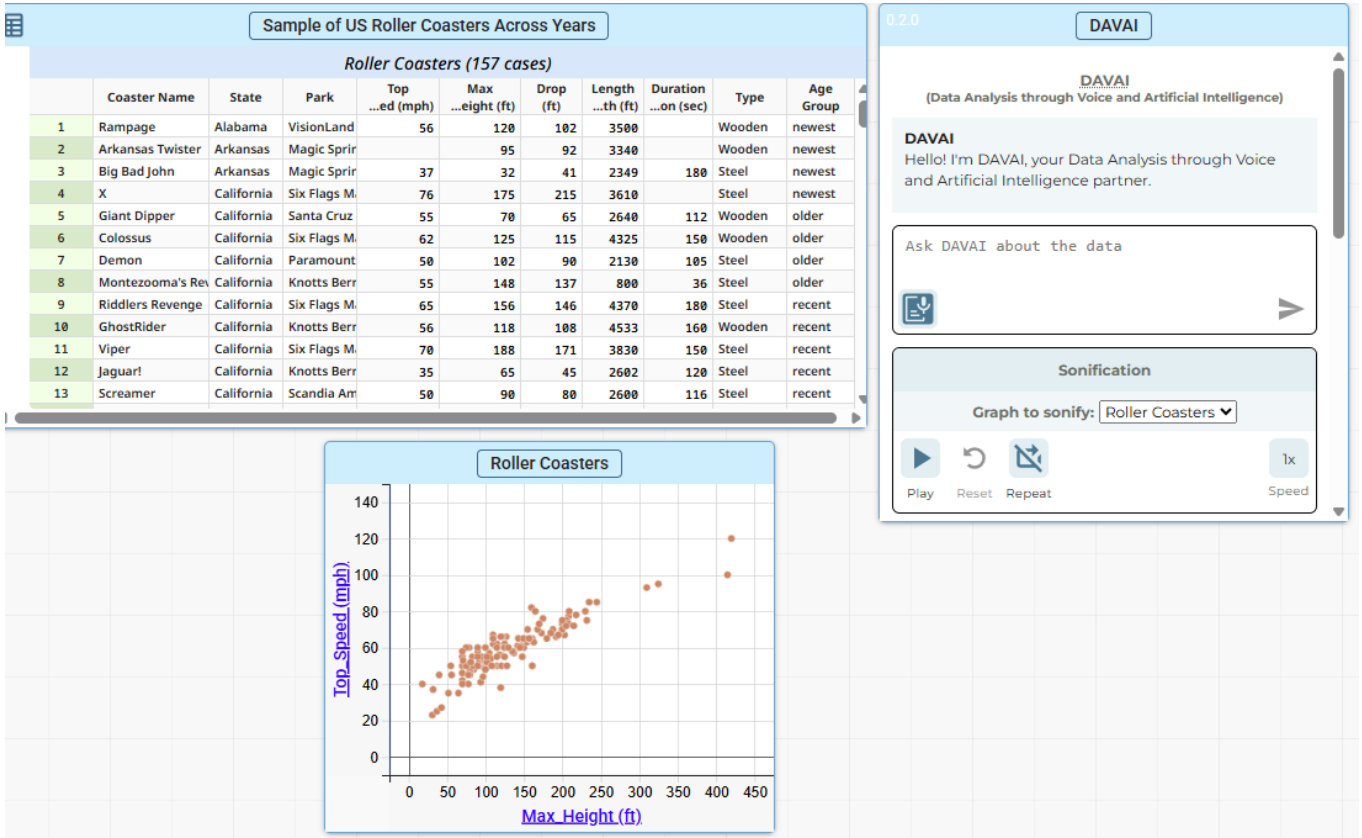


The Concord Consortium

Current Implementation



CODAP/DAVAI Demonstration



The Concord Consortium

Challenges

- Finding a sufficient pool of testers
  - Locating science/math blind/LV students who can attend in-person testing sessions is more challenging than expected
- Speed: DAVAI can take a long time to process requests
  - We are reconfiguring how DAVAI processes requests to reduce response time
- Accuracy of agent actions: Doesn’t always do what is intended
- Consistency of response format: DAVAI’s responses sometimes contain irrelevant information (e.g., hex codes of dot colors)

Testing: Observations

- Students seem to favor asking DAVAI for general data descriptions, rather than asking it for specific details
- Students in general and possibly BLV students in particular don’t have a lot of skill in independent exploration of data
- Each student tends to interact with DAVAI in their own way with a variety of access tools and idiosyncratic language (range of vocabulary based on past experience)
- Everyone testing so far has indicated that DAVAI/CODAP has been much easier than trying to work with data in spreadsheets

Testing: Feedback/Improvements (1)

- DAVAI has the potential to help blind/low-vision students explore data and graphs
- DAVAI's descriptions are generally useful
  - Sometimes they are excessively verbose
- Sonification is very useful, especially the sound moving from left to right as it plays
  - Add baseline/top reference sounds to establish range context
  - Ensure gaps are represented in sonification
  - Allow users to add markers to the sonified graphs

Testing: Feedback/Improvements (2)

- When DAVAI responds, start with a summary, then provide details
- When describing data, DAVAI should include the minimum and maximum values as a frame of reference
- Add support for more graph types (currently limited to scatter plots and single dot plots)
- Add a sonification shortcut command as well as speed controls
- Add a sound that indicates the beginning of the sonification (useful when looping)
- Add a note-taking tool or scratchpad
- Add a read-aloud function
- Add a text-to-speech function, with controls for pause/play, speed, stepping forward/backward

Technology RQ: In what ways can generative AI/LLM-based technologies or approaches be extended, designed, and leveraged to facilitate accessible interaction with data for BLV users?

Learning RQ: What effect does the availability of interactive and generative technologies have on BLV students' ability to engage with and make meaning of datasets?

Questions?

The Concord Consortium

- Daniel Damelin - PI - ddamelin@concord.org
- Bill Finzer - Senior Scientist - wfinzer@concord.org
- Lynn Stephens - Researcher - lstephens@concord.org
- Cynthia McIntyre - PM - cmcintyre@concord.org

onfocusAccess

- Geoff Freed - Co-PI: geoff.freed@onfocusaccess.org