@CONCORD

vol. 28 • no. 2 • Winter 2025

The Concord Consortium
concord.org

Perspective:
Historic Innovations in Educational Technology

By Chad Dorsey

The Concord Consortium has just celebrated its thirtieth anniversary. As I flip through the first issue of our *@Concord* newsletter, I feel both nostalgic and proud. That original issue paints a picture of an idealistic and scrappy era. Across its articles, I see innovators grappling with technical hurdles—that perennial hallmark of cutting-edge work—as they struggle to ensure broad access to technology-based resources in an era when computers were still hard to come by. I also see remarkable examples of ingenious ideas becoming reality, in the form of theories that would come to enter the common vernacular and early-stage software concepts that would ultimately transform into entire industries. From online learning to sensors and simulations, the Concord Consortium has been pioneering educational innovations from its inception.

The historical artifacts of our founding provide the closest thing we can get to a crystal ball. If we take a similar snapshot of technology today, what might it portend for the teaching and learning landscape of tomorrow? Let's take a look back as we try to envision the future.

An early newsletter article* provides a detailed account of the circuitous technology path involved in learning how to support teacher and student interactions in online courses. A series of vivid descriptions of software packages tried and discarded make evident the vital role of a custom, integrated software environment in supporting asynchronous online learning. The article outlines a set of discoveries: the need for threaded discussions, the importance of sharing resources that include images, and the ability for participants with little technical background to easily create new discussions. In hindsight, it is clear that those experiments to make virtual education possible across time and space marked some of the earliest concepts of a full-featured K-12 learning management system.

Another article outlines the emergence of a broadly accessible computer for classroom use. Descriptions of the Apple eMate frame it as ideal for the K-12 learning setting and "far more valuable educationally" than the more powerful—and more expensive—office computers of the time. Affordability and usability were the criteria used to nominate that early portable device, calling to mind the introduction and meteoric rise of Chromebooks as they later came to dominate the K-12 educational market.

Other prophetic examples abound. One article makes the case for "collaborative creation of graphics," predicting that "maps [of environmental study sites] with overlays showing…features and collection sites" could change the way students think about science. Students were able to interpret and analyze data from far-off places. Other articles outline how "inexpensive, portable computers…could transform education," predicting that "lifelong learning…can be made universally available."

Indeed, these visions of online learning from three decades ago have become reality. Today's ubiquitous devices and networks enable access in any setting imaginable. And YouTube's endless catalog of tutorials on practically every aspect of the human experience supports all manner of informal learning as well. It's now possible to learn about anything, anywhere and anytime. It's also clear that thanks to the firm establishment of virtual learning across the nation, schools were able to go remote during the pandemic. While we are still unpacking the effects of that grand experiment, technology was central in connecting people across an incredible diversity of settings.

Looking forward

So what does this historical viewpoint offer us today? For one, it reminds us that our work in technology innovation leads to meaningful future developments. While some innovations will surely prove to be first of their kind, others will provide stepping stones or incremental improvements. But which elements of today's initiatives presage new domains or essential future affordances?

Chad Dorsey
(cdorsey@concord.org)
is President and CEO of the Concord Consortium.

Our models enable students to investigate these systems via inquiry-based experiences, allowing them to change parameters, predict outcomes, and reason about events and processes.

What seemingly small innovations are waiting to evolve into whole new industries ten or twenty years from now? While the only thing certain about all predictions is that some will be wrong, speculation is instructive, nonetheless.

It is already clear that our work promoting quality Earth science learning is groundbreaking. Over the past decade, researchers at the Concord Consortium have developed a suite of dynamic models for understanding geology and Earth systems. Because of the long-term nature of geological time, Earth science largely involves the study of static situations and disconnected ideas. The Earth, however, is a highly dynamic, connected system, and its rocks exist in different forms for key reasons—reasons that connect to the Earth's system-focused nature. Our models enable students to investigate these systems via inquiry-based experiences, allowing them to change parameters, predict outcomes, and reason about events and processes.

These authentic, complex Earth and environmental system models empower students to be "hands on," exposing them to simulation as a tool that can open doors for investigating physical systems and drawing important conclusions. The same goes, of course, for other subjects—models and simulations play a key role in examining science both inside and outside K-12 schooling.

In other areas of STEM education, we are integrating software programming and hardware devices to enable easy programming and deployment of microprocessors and Internet of Things-like devices. This infrastructure may well evolve into an essential paradigm for guiding and launching hardware. Indeed, as robotics emerges in ways difficult to predict, this future seems more plausible every day.

We also are enabling students to envision mathematical modeling in a flexible environment that links elements from text to drawings and equations. While mathematical modeling has long been important, it has seen a resurgence in recent years within mathematics education circles and even at the level of statewide programs and standards. Combining this increased awareness with the rising relevance of modeling in the context of machine learning and AI-based approaches, it's not difficult to imagine this work as a template for indispensable learning tools in future classrooms.

Finally, the rapidly developing world of AI is sure to create persistent traces for future generations of educational technology. At the Concord Consortium we are incorporating AI technology into our collaborative learning platforms to help teachers quickly and easily interpret, categorize, and react to students' sensemaking. We are also working to help students themselves learn about the technologies underlying AI systems, preparing them to understand the tools that are rapidly shaping their future so they can approach these tools from an informed perspective. There is little doubt that this critical work is destined to evolve in essential ways for teachers and students alike.

These are only a few examples of how the present can become the future. Ultimately, we hope all our work will prove consequential in both its direct effects as well as the distant consequences our projects have for innovations and inspiration for equitable, large-scale improvements in STEM teaching and learning through technology. We remain as hopeful about this ambitious goal as we were back in that first issue of *@Concord*. Were those early glimmers of technology to succeed, our inaugural newsletter predicted, "far more people can share their knowledge with many more interested learners." That goal is as relevant today as it was then, and we aim to continue this important work for many more decades.

* Read our first newsletter at https://concord.org/newsletter/1997-spring/

Data Science, AI, and You in Healthcare: Teaching Students About Medical Data Bias

By Kathy Jessen Eller

Kathy Jessen Eller
(kjesseneller@concord.org)
is a research scientist.

Data science and artificial intelligence (AI) are poised to upend the way healthcare is delivered. But what happens when AI is trained on medical data that doesn't reflect the population? A new semester-long course teaches high school students how embedded bias in machine learning affects healthcare discussions and decisions. The goals of the Data Science, AI, and You in Healthcare project include fostering community connections between educators, researchers, clinicians, and local stakeholders to prepare underrepresented students for STEM jobs—ultimately in fields where their background and experience can help mitigate bias embedded within AI healthcare models and improve outcomes for all.

Designed for students interested in AI, medicine, and social justice, the Data Science, AI, and You (DSAIY) in Healthcare course has no statistics or computer science prerequisites. Students are introduced to data science practices through several engaging activities, such as "What's Going On in This Graph?" in *The New York Times*, and by visualizing data from non-medical current events that interest students, for instance, data from Taylor Swift's concerts. They use CODAP to further explore and visualize data, including descriptive box plots and other types of graphs.

DSAIY also includes fun hands-on activities designed to make machine learning concepts more accessible, for example, using candy to classify and split a dataset into two groups, one used to train data and the other to test the model. Students also interview family members about AI, create posters of under-recognized scientists of color, and hear directly from a community member invited to the classroom to share stories about their personal experiences with medical bias. Finally, the course concludes with a collaborative datathon.

The DSAIY course has undergone three revisions since it was developed in spring 2023 via iterative feedback from teachers, students, and an interdisciplinary team within the program-created Hive Learning Ecosystem. This ecosystem emphasizes cross-disciplinary multigenerational collaboration and includes high school students, teachers, undergraduate mentors who help in the classrooms, data scientists (graduate students, postdoctoral fellows, and tenured professors), and members of the medical community across the education pipeline (undergraduates, medical school students, residents, fellows, and clinicians).

Background

During the height of the pandemic in 2021, a group of educators, data scientists, and clinicians got together and asked, "How can we increase the diversity of people entering data science, AI, and healthcare to mitigate bias organically?" Our idea was that the more diverse we make the fields, the more people will recognize, call out, and mitigate factors leading to biased predictions.

To get a sense of what we mean by biased predictions, take a look at Figure 1. It shows three screenshots of the Medical Priority App used in a code.org unit on AI and machine learning and trained with data from real patients in 2020. As part of the DSAIY course, students simulate signing in to the app as patients. They enter age, race, and gender, as well as urgent medical concerns, such as whether they are bleeding, experiencing shortness of breath, where they are feeling pain, and what type of pain it is—data that would typically help a medical professional triage the patient.

Compare Figure 1A to 1B. Notice that all entries are the same except for gender. If you're a White (W) male, you are a priority. But if you are a White female, your status is normal, in other words, not prioritized for care. Now compare 1A to 1C. Both are male, but one is White and the other is a Native Hawaiian and Pacific Islander (NHPI). The White male is a priority while the NHPI individual receives a note to "return later." Through the DSAIY course, students ask questions about how the bias came about, whether they think it was intentional, and how the training data might have affected the machine learning algorithm.

Here's another example. The pulse oximeter was invented in 1974 to measure blood oxygenation level quickly and easily by reading

A B C

Figure 1. Screenshots from code.org's Medical Priority App with various sample patient scenarios.

the amount of light absorbed when passed through the skin. The amount is proportional to the concentration of hemoglobin (oxygenated blood) in the blood vessel. Low light absorption means a relatively low amount of oxygenated blood. An algorithm built into the pulse oximeter translates this information into a percentage of oxygenated blood. Anything below 94% is of concern. Unfortunately, research shows pulse oximeters are prone to error when there is more melanin in the skin. The more melanin, the more likely the reading will erroneously default to 100%. Thus, many people with more pigmented skin, particularly Black people, could be turned away from treatment because it appears as though there is plenty of oxygen in their system.

These false readings had critical implications during the COVID-19 pandemic.[1] Imagine hospital waiting rooms during the initial outbreak. They were overflowing with people coughing and having trouble breathing. Nurses quickly measured vitals with the pulse oximeter and other devices. Yet many people who were struggling to breathe and had dangerously low arterial oxygen levels were turned away. Why? Because the pulse oximeter measurement—with its embedded bias—presumed that all was well.

Datathons

At the end of the Data Science, AI, and You in Healthcare course, students participate in a two-day datathon. Teams of at least two high school students, one STEM teacher, an undergraduate computer science student, a clinician, and a data scientist collaborate to solve a problem. The cross-disciplinary collaboration is a critical part of the DSAIY experience, as described by a teacher:

> *"One of the things I loved about the datathon was the students seeing cross-disciplinary teams working together on a common goal. Seeing doctors who admitted they didn't know anything about datasets working with data scientists who were asking questions about the pulse oximeter. I think that was really powerful for kids to see."*

The datathons are based on a model developed by DSAIY Co-Principal Investigator Dr. Leo Celi, a senior research scientist at Massachusetts Institute of Technology and a staff physician at Beth Israel Deaconess Medical Center, to encourage diverse perspectives to collaborate around healthcare problems and warn about unintended consequences that have the potential for causing harm. At the 2023 MIT Critical Care Datathon, teams worked to create a pulse oximeter correction factor. Although this was beyond the scope of the two-day event, participating mentors later published a correction factor from this work.[2] During the 2024 Brown Health AI Systems Thinking for Equity Datathon, teams used Women in Data Science mortality training data to show bias in their own machine learning models. In addition, they read and discussed articles detailing how large language models could be queried to provide personal information, even if the models were not trained to do so.

Next steps

We have piloted the DSAIY course in Rhode Island with over 300 students and 12 teachers who significantly improved the curriculum with each iteration. One teacher reported a student "talking about how he knew he wanted to do computer science, but he was really thinking about going into the data sciences now because 'data is everything.'" Analysis of student pre- and post-surveys confirms that students learned about data science, AI, and related careers by participating in the course.

1. Personal communication from Dr. Leo Celi.

2. Matos, J., Struja, T., Gallifant, J., Charpignon, M. L., Cardoso, J. S., & Celi, L. A. (2023). Shining light on dark skin: Pulse oximetry correction models. In *2023 IEEE 7th Portuguese Meeting on Bioengineering* (ENBENG), 211–214.

LINKS

Data Science, AI, and You in Healthcare—
concord.org/dsaiy and www.dsaihealthed.org/

(continued on p.6)

(continued from p.5)

Kathy Jessen Eller is a research scientist at the Concord Consortium. **Dr. Leo Celi** is a senior research scientist at Massachusetts Institute of Technology and a staff physician at Beth Israel Deaconess Medical Center.

Q. What drew you to the topic of bias in healthcare?

Kathy: I believe it's a human right for a person to have healthcare. It didn't hit me how much healthcare is a privilege in the U.S. until I learned more about poverty and socioeconomic struggles when I taught middle school in Providence, RI. I met Dr. Celi through a mutual colleague because we were interested in writing a proposal for a new NSF initiative in equity and STEM. He introduced me to the bias in healthcare.

Leo: My Eureka moment came when I read a paper saying that if you look at the pictures in medical textbooks, they're all White people. It made me realize that science is created in a non-neutral space. Science is created by people who have their own experiences, their own motivations. That grew my mistrust of everything that I use to practice medicine. So I thought, maybe we can use all this data that is collected and create a much more inclusive knowledge system. AI is truly exposing the breadth and the depth of biases in the way we deliver care and if we're not careful this will all be encrypted in algorithms.

Q. What excites you about the DSAIY curriculum?

Kathy: The authenticity of the program excites both me and the students. The program imbues computer science with real data and includes real people who share their experiences with bias in healthcare. The first time we taught this course there was a young mother at the Met High School in Providence whose own son was sent home because his pulse oximetry reading was read as positive. She had to bring him back to the ER because he was struggling to breathe.

Leo: What excites me is that this project is instilling critical thinking at a young age. Critical thinking overlaps with systems thinking, which overlaps with data science. The way you evaluate systems is by looking at their data exhaust. And by doing that, you also realize that there's so much in the data that's missing and that to be able to really redesign systems, we might need to collect more data.

Q. How do you hope students will be affected by this experience?

Kathy: We hope to increase students' awareness that healthcare inequities exist and empower them to advocate for themselves, their family members, and communities. We also hope that the course will increase students' awareness of the types of careers available in healthcare and support their STEM identity so that they will think of themselves as being capable of a career in a position of power in data science and healthcare.

Leo: I'm hoping that DSAIY enriches their lives in a way that they had no idea before they enrolled in the program and that they will take advantage of the networks that they get exposed to. I've continued to work with high schoolers who are now applying to grad school, and what they say they benefit from and value is how the experience exposed them to a new world.

Q. What does the datathon add to the DSAIY curriculum?

Kathy: Mentors, including data scientists, medical professionals, and undergraduate students volunteer at the datathon to encourage students to go into related careers and teach them about how much impact bias can have on machine learning predictions. It's very empowering for students to hear a medical doctor ask a data scientist questions and vice versa. No one knows everything. Collaboration is the key to solving complex problems, like the impact of bias on healthcare.

Leo: The datathon allows us to think with people who don't think like us. The way to redesign the system [to mitigate bias in healthcare discussions and decisions] is to bring in people who don't think like us. That includes computer scientists and social scientists. They have not been involved in creating and validating medical knowledge. And I think their perspective, which is very different from ours, will be the secret sauce. High school students also give me a fresh perspective because I don't think like they do.

Data Science, AI, and You in Healthcare

Monday's Lesson:

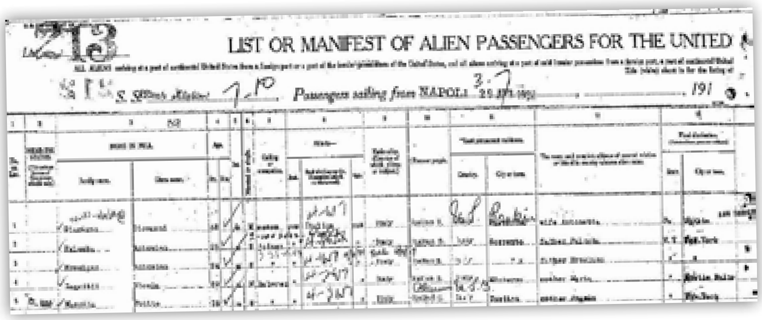
Immigration Data and the Great American Melting Pot

By Rachel Folger and Kate Miller

Rachel Folger
(RFolger@schools.nyc.gov)
is a middle school social studies teacher in New York City.

Kate Miller
(kmiller@concord.org)
is a research associate.

Middle school students typically study the immigration experience of the tumultuous 20th century and its impact on American policy in their social studies class, often using photos and written accounts as primary sources. Data can also serve as an engaging primary source. Data science is an inherently interdisciplinary field, allowing students to make connections between humanities, math, and issues of community interest.


LIST OR MANIFEST OF ALIEN PASSENGERS FOR THE UNITED

Passengers sailing from NAPOLI ... 191

Figure 1. Ship manifest from the National Archives.

This lesson is useful when studying U.S. immigration policy in the late 19th and early 20th centuries, the reasons for immigration, and the stories of immigrant experiences at both Ellis Island (New York) and Angel Island (San Francisco). Consider including discussions around the Chinese Exclusion Act and the 1924 Immigration Act. While you can start this lesson on a Monday, plan for three to four class periods to finish.

The hook

Show your students an image of a ship manifest from an arrival at the Port of New York (Figure 1). Then guide students through a conversation about what counts as data in social studies. (Students might consider, for example, immigrants' arrival date, country of origin, or nationality.)

Introduce the data

Now show a screenshot of the same information from a ship manifest, this time displayed as a table in CODAP (Figure 2). Ask students: What do you notice? What do you wonder?

As students share what they notice and wonder about the manifest, guide their attention to the column titles. These are the data attributes. Additionally, if not mentioned by a student, point out that some people traveled beyond the U.S. (their final destination is outside the U.S.). To encourage students to wonder about different travel experiences, choose one data case (row) and ask what it might have been like to be that person. How old were they? Were they traveling alone or with their family?

Explore immigration trends

Working in pairs, students should open the Ship Logs Sample CODAP document (http://short.concord.org/m0x) and explore the dataset. The goal is to identify similarities or differences between people arriving in New York City and those arriving in San Francisco. For example, click the Graph button in CODAP, then drag the Arrival Port attribute to the y axis and the Arrival Date to the x axis. Continue making additional graphs. Note: Students can save their CODAP documents to Google Drive or as a local file.

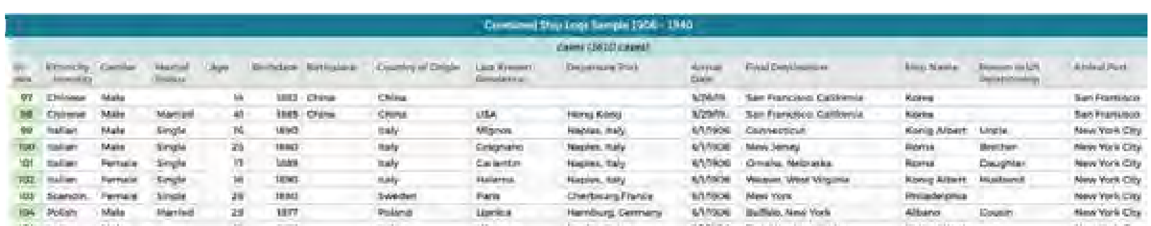
Ask students: What did you learn from graphing these attributes? Does the graph show a similarity or difference between immigration to New York and San Francisco? Why do you think a similarity or difference exists?

Tell an immigration data story

Students should use their data exploration to support a story about immigration. Click the Plugins menu in CODAP and choose Story Builder, then click the question mark in the upper right for a short tutorial. (Alternatively, students can use Google Slides or another presentation format.) Students' data stories should be supported by graphs from the dataset and include an answer to one or more prompts:

- Who immigrated and to where?
- Why did people immigrate? Was there a difference between Asian and European immigration?
- When did people immigrate? How did laws and world events impact immigration?

Through data storytelling, students can enhance their critical thinking, research, and analytical skills in humanities classrooms and beyond.


Combined Ship Logs Sample 1906 – 1940

cases (3610 cases)

index	Ethnicity	Gender	Marital Status	Age	Birthdate	Birthplace	Country of Origin	Last Known Residence	Departure Port	Arrival Date	Final Destination	Ship Name	[illegible]	Arrival Port
97	Chinese	Male		[illegible]	[illegible]	China	China			[illegible]	San Francisco, California	Korea		San Francisco
98	Chinese	Male	Married	41	1865	China	China	USA	Hong Kong	[illegible]	San Francisco, California	Korea		San Francisco
99	Italian	Male	Single	16	1890		Italy	[illegible]	Naples, Italy	6/1/1906	Connecticut	Konig Albert	Uncle	New York City
100	Italian	Male	Single	25	[illegible]		Italy	[illegible]	Naples, Italy	6/1/1906	New Jersey	Roma	Brother	New York City
101	Italian	Female	Single	17	1889		Italy	[illegible]	Naples, Italy	6/1/1906	Omaha, Nebraska	Roma	Daughter	New York City
102	Italian	Female	Single	16	1890		Italy	Palermo	Naples, Italy	6/1/1906	[illegible], West Virginia	Konig Albert	Husband	New York City
103	Scandin.	Female	Single	26	1880		Sweden	Paris	Cherbourg, France	6/1/1906	New York	Philadelphia		New York City
104	Polish	Male	Married	29	1877		Poland	Lipnica	Hamburg, Germany	6/1/1906	Buffalo, New York	Albano	Cousin	New York City
105	Italian	Male		15	1891		Italy	Udine	Naples, Italy	6/1/1906	Brooklyn, New York	Konig Albert		New York City

Figure 2. A CODAP data table from combined ship logs (1906–1940).

LINKS

CODAP—codap.concord.org

Trudi Lord
(tlord@concord.org)
is a senior project manager.

BUILDING STEM Identity Through YOUTHQUAKE

By Trudi Lord

Consider for a moment how you think about yourself in relation to science, technology, engineering, or math. If you're a STEM teacher, instructional designer, or researcher, you may identify as someone who knows about, cares about, and contributes to STEM pursuits. You may have developed these feelings as a child who was encouraged to question how things worked, just as engineers do. On the other hand, you might have originally shied away from math and science, thinking that those subjects were for other people, but not for you.

Research suggests that strengthening STEM identities can help students who are traditionally underrepresented in STEM, including girls and disadvantaged students, feel like they belong in these fields. The YouthQuake project brings authentic learning experiences in computational geoscience to students at a critical point in the development of their STEM identity—middle school.

Starting close to home

Central to developing STEM identity is building on students' own experiences. Most students in California live within 30 miles of an active seismic fault. Many have experienced, or know someone who has been affected by, an earthquake. The YouthQuake project team and teachers have co-designed a curriculum unit about earthquakes from both scientific and community perspectives. One focus is on the many careers related to earthquakes—from geoscientists who track GPS land movement to firefighters who help communities prepare for and respond to natural disasters.

At the core of the YouthQuake curriculum is a five-activity online module that guides students through a series of increasingly complex investigations of earthquake risks and hazards in California. Students are introduced to block-based coding using the GeoCoder and real-world GPS data to investigate how the land in and around the state moves and deforms (Figure 1). Students also examine the locations of plate boundaries along the West Coast and investigate earthquake epicenter data using Seismic Explorer (Figure 2). By giving students the opportunity to use authentic geoscience tools and practices to explore the land movement and earthquake history of their community, we hope to foster students' awareness of and interest in computational geoscience careers.

Role models

Equally important to using the tools of the trade is meeting practicing scientists, including people of the same gender and racial identities who can serve as role models. Through the YouthQuake

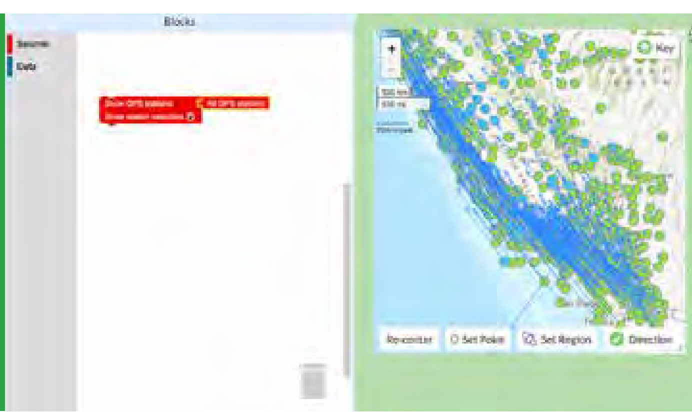


Figure 1. GeoCoder allows students to explore GPS motion data in California. Students add simple code blocks (left) and observe the output on the simulation (right).

Figure 2. Students investigate earthquakes with Seismic Explorer.

Figure 3. The project team tests a snappy hands-on activity about deformation using rubber bands before adding this activity to the YouthQuake curriculum.

Figure 4. Small groups of teachers and curriculum developers share potential STEM identity-strengthening activities.

curriculum, students were introduced to early-career geoscientists, who Zoomed into their classrooms and answered questions about their jobs, school experiences, and career aspirations. The scientists—four women and one man from diverse cultural and racial backgrounds—described taking different paths to their STEM careers. One African American scientist told the students that she wished she had had a scientist talk to her class when she was in middle school because she didn't know that there were jobs in science for people who looked like her. After the virtual visit, a student told her teacher, "I didn't know women could have these kinds of jobs."

Co-designing additional STEM identity activities

Following the implementation of the module in spring 2025, we held a summer institute with the three pilot YouthQuake teachers and three new teachers, focusing on new hands-on science activities (Figure 3) and additional approaches for nurturing students' STEM identities (Figure 4). We began by looking at the results of our STEM identity survey, designed to measure students' attitudes and thoughts about STEM as well as their knowledge of STEM careers (specifically in science, Earth science, and computer programming) and practices.

Many students showed significant improvements in their attitudes toward Earth science, knowledge of earthquakes and scientists' practices, and their thoughts about "people like me" in Earth science. While our survey results were generally positive, it was clear that we could do more to build students' STEM identities and make connections to future careers.

The teachers described how the YouthQuake curriculum connected to students' developing STEM identities. For example, the block coding appealed to some students who had previously participated in a robotics afterschool program and thought of themselves as "coders." For other students, including the girl described above, the virtual visits from young scientists were the most transformative. Finally, other students were most engaged by an in-person presentation from firefighters about earthquake safety and preparedness.

The curriculum co-design team of pilot teachers and project partners then looked inward, thinking about our own STEM identities. We paired up and shared our own personal journeys into STEM careers. As self-aware adults, we were able to look back and describe how our identities have grown and been shaped over time.

Recognizing that middle school students are just beginning to build their identities, teachers felt that they needed to directly address STEM identity with their students. Students, they argued, should have the opportunity to reflect on their STEM identity and explore the next steps in their STEM identity formation, as well as the multiple pathways that could lead to a range of STEM careers.

A sense of belonging in STEM

Beginning early in the school year, these science teachers plan to explicitly discuss STEM identity with students, asking the not-so-simple question, "What is an identity?" They expect to return to this question throughout the year and to have students construct roadmaps to different careers, including researching the necessary coursework or other training to reach that career.

During the YouthQuake module, teachers plan to prompt students to think broadly about the different jobs that go into everything from installing GPS stations to monitoring earthquake activity to building shake-proof buildings. Teachers also plan to integrate STEM identity into the YouthQuake end-of unit-project.

To foster curiosity about STEM careers, students will each take on the role of an "expert" on a team of geoscientists, selecting a traditional STEM or STEM-adjacent job involving earthquake science, preparedness, or recovery—from a seismologist to a computer programmer or EMT. Students will research the career and career path and present what they learned from the YouthQuake curriculum to parents at the school science night or to students in younger grades.

Through such opportunities, we hope that students stop thinking "science is not for people like me" and begin to align their current identity with who they want to be. A strong STEM identity is more likely to result in continued coursework in STEM subjects and eventually the pursuit of STEM-related careers. Importantly, more diverse perspectives can lead to more innovation solutions, and in this case, better understanding of and preparation for natural hazards that California may face in the future.

LINKS

YouthQuake—concord.org/youthquake

Using Sound to Enhance Data Interpretation

Dan Damelin
(ddamelin@concord.org)
is a senior scientist.

By Dan Damelin

The use of sound to communicate nonverbally predates written records and is deeply embedded in human history. The simple rhythmic sounds from early drums, bells, whistles, and horns evolved to become complex communication data. Today, graphs are the most common way information from data is conveyed. The COVID-Inspired Data Science Education through Epidemiology (CIDSEE) project, a collaboration between Tumblehome, Inc. and the Concord Consortium, is developing and researching new tools that add sound to graphs to help students explore and make sense of data.

Sonification is the process of turning information into sound, for example, to mark time with bells or relay battlefield information with drums or trumpets. It also enhances scientific discovery by providing a multimodal approach to data analysis, communicates meaning through artistic interpretations of data, and increases the accessibility of data visualizations.

A short history of sonification

One of the first modern examples of sound being used to represent numerical data occurred in the 1920s when geoscientists turned seismic data into sound, allowing researchers to "hear" earthquake patterns. (Seismographs themselves resemble sound waves [Figure 1].) In the 1930s and 1940s, auditory signals provided situational awareness for pilots and navigators, including information about the condition of their planes and ships, as well as proximity warnings.

From the 1950s to the 1970s, even as computers were just being developed, sound synthesis and a rudimentary exploration of representing data using sound laid the foundation of modern sonification. Bell Labs developed one of the first programs for digital sound synthesis. While primarily focused on creating music, this work demonstrated how numerical data could be translated into sound, inspiring collaborations between artists and scientists who explored data-driven composition and influencing later scientific sonification. About the same time, NASA researchers also began experimenting with mapping data to sound to interpret large datasets collected during space missions. Such sonification allowed scientists to explore patterns in data that were difficult to discern visually.

In 1992, Gregory Kramer organized the first International Conference on Auditory Display, marking a major milestone in the history of sonification. The event brought together researchers from diverse disciplines to discuss sonification and formalized sonification as an academic field. Kramer and others developed frameworks for how sonification could be systematically applied to scientific and practical contexts. The International Community for Auditory Display, which began as a result of that inaugural conference, is still active today.

In recent years, as computers have become more powerful, sonification has become more ubiquitous. It has been used across multiple disciplines and applications from EEG monitoring to sifting through the enormous datasets generated by the CERN particle collider to hearing trends in climate data.

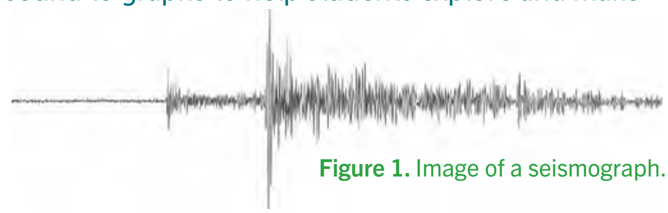


Figure 1. Image of a seismograph.

CODAP and the Sonify plugin

The Common Online Data Analysis Platform (CODAP) is a free, web-based tool for supporting students in sensemaking with data and learning data science practices. CODAP's functionality can be extended through plugins. In 2018, Takahiko Tsuchiya, then a graduate student at the Georgia Institute of Technology, began building a plugin with the help of one of the CODAP developers, Jonathan Sandoe, to facilitate sonification of data in CODAP. The National Science Foundation–funded CIDSEE project has extended and improved the Sonify plugin, which is available from the plugin menu on the CODAP toolbar.

Figure 2 shows an example of the Sonify plugin, configured to work with a dataset of CO_2 measurements over time from NOAA's Mauna Loa observatory. The graph of that data shows a classic upward trend (Figure 3). To sonify this data, the user selects what attribute to map to pitch and what attribute to map to date/time. After the user clicks "create graph," the plugin instructs CODAP to generate the graph of CO_2 vs. date. When the user clicks "Play," the green line on the graph sweeps from left to right and the plugin produces a sonic representation of the data with higher values of CO_2 generating higher-pitched tones.

Figure 2. The Sonify plugin in CODAP.

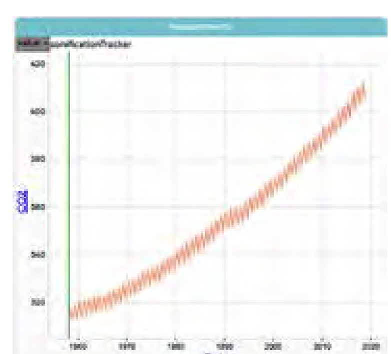


Figure 3. Graph of CO_2 vs. date at the Mauna Loa observatory in Hawaii.

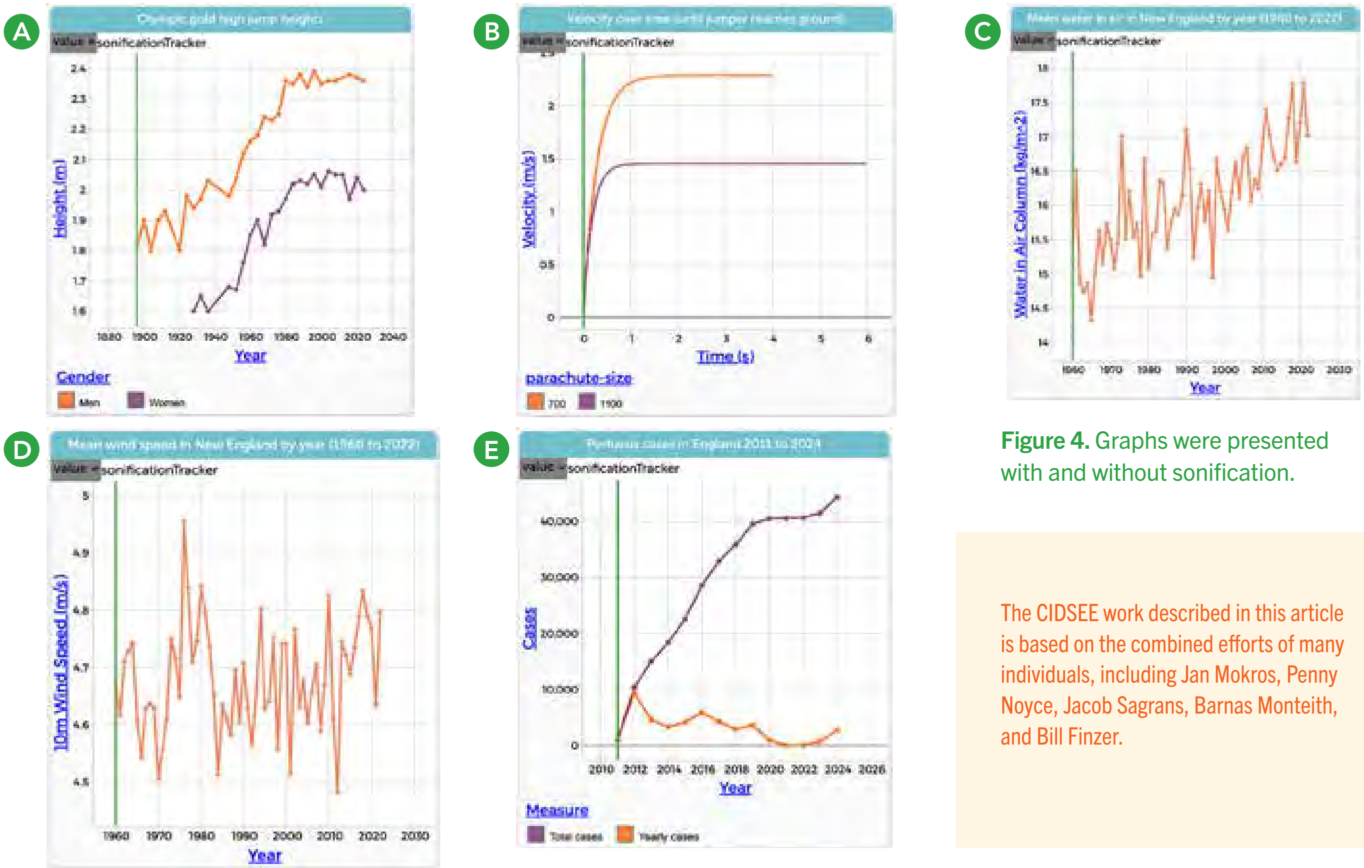


Figure 4. Graphs were presented with and without sonification.

The CIDSEE work described in this article is based on the combined efforts of many individuals, including Jan Mokros, Penny Noyce, Jacob Sagrans, Barnas Monteith, and Bill Finzer.

CIDSEE and sound graphs

Since 2021, the CIDSEE project has empowered over 1,000 underserved middle school students in afterschool and summer camp "Data Detectives Clubs" to explore the field of epidemiology through the use of scientific data and innovative data tools. By engaging with real-world epidemiological challenges, students cultivate critical skills and an appreciation for the importance of data in understanding public health issues. The program is designed to enhance students' confidence, interest, knowledge, and abilities in working with datasets and data tools, including CODAP and its Sonify plugin, to investigate pandemics. The goal is to inspire student interest in epidemiology and related STEM fields, including research, modeling, data analysis, and science communication, by providing hands-on experiences with the data-driven work of epidemiologists.

Through a supplemental award from NSF, the CIDSEE project is extending its work to study how and when sonification might aid student understanding of time-series graphs, such as the spread of COVID over time, and improve data interpretation with other time-series phenomena. To explore the value of sonification, we asked students to interpret a variety of graphs with and without sound (Figure 4, A-E):

A) a comparison between two groups (e.g., men's vs. women's winning Olympic high jump heights over the years)

B) a simple phenomenon with an easy description (e.g., it goes up then levels off)

C) a clear trend with some significant variability (e.g., water content in the air over time)

D) no clear trend (e.g., change in wind speed over time)

E) a cumulative vs. marginal comparison (e.g., yearly vs. cumulative pertussis cases)

While the analysis of the study is ongoing, some preliminary observations have emerged:

- Most students easily grasp the basic concept of sonification, understanding that higher pitches correspond to higher values and lower pitches to lower values.
- Students appreciate having the ability to adjust the playback speed, allowing them to listen to the graph at a faster or slower pace.
- Overall, students find sonification appealing—it's a novel and engaging way to interact with data.
- Students find it useful to have the visual line scan across the graph during sonification, helping them track their position on the graph while listening.
- Most students find it challenging to interpret the sonification through sound alone. They need to see the graph alongside the audio to make meaningful sense of the data.

Accessibility and sonification

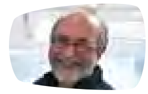
A new NSF-funded project called Data By Voice is building upon lessons learned by the CIDSEE project to explore ways to make data sonification more understandable, especially for blind and low-vision (BLV) users. The project is using artificial intelligence to facilitate sensemaking with data in CODAP by creating a voice interface for interacting with CODAP and a large language model to describe data representations created by BLV students. In addition to verbal descriptions generated by AI, students can request that a graph they made be sonified. With AI, the future of sonification is just beginning.

LINKS

CODAP—codap.concord.org
CIDSEE—tumblehome.org/research/
Data By Voice—concord.org/data-by-voice

Alaskan Students Profile Their Local BEACH

Joy Massicotte
(injoy5683@gmail.com)
is a curriculum designer
and consultant.



Tom Moher
(moher@uic.edu) is an Emeritus
Associate Professor at the
University of Illinois at Chicago.

Carolyn Staudt
(cstaudt@concord.org)
is a senior scientist.

By Joy Massicotte, Tom Moher, and Carolyn Staudt

Beaches are dynamic landscapes that are always changing. However, due to human development and climate change, the speed and intensity of change has been increasing. Our Precipitating Change with Alaskan Schools project has developed curriculum activities focused on the coast and coastline so students can engage in locally relevant scientific practices.

Alaska has over 6,600 miles of coastline, or nearly 34,000 miles if you include islands and their shorelines as well. By any measure, that's a lot. The coastline is critical to Alaska's Indigenous students and their families who rely on the ocean for subsistence practices of hunting, fishing, and gathering traditional foods. The goal of our Precipitating Change project is to provide opportunities for students to learn more about the coast using both Indigenous and Western science approaches and to share their new understandings and stories with their communities.

We collaborated with teachers to create an alternative format for activities in our online learning management system. We wanted to offer students a more open-ended learning experience to record their observations rather than a linear sequence of online activity pages and typical multiple-choice questions. Designed to look like a spiral notebook with tabs for sections, the new online logbook invites student-led inquiry (Figure 1). Students can upload images, create drawings, and type their notes, including through audio recordings. The logbook also contains traditional curriculum activities, such as videos (e.g., of Elder interviews and coastal protection strategies), models of wave motion such as during a storm surge, real-world data, an interactive glossary, and more.

Figure 1. The logbook format was co-designed with teachers to be more student driven and to include more than the traditional question and answer format.

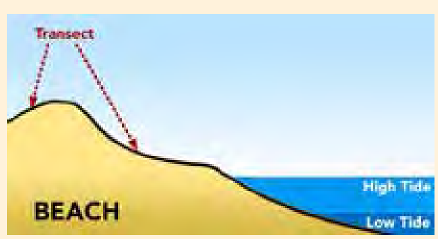


Figure 2. A transect is a cross-section view of the beach.

Throughout the logbook, students are prompted with special "think about…" sections. For example, students are asked to think about their experiences harvesting food from the ocean; the impact of climate change on the food supply for their village; stories from parents, community members, and Elders about changes to the beach; and adaptations they would recommend for their beach. In this hands-on, place-based curriculum, students are encouraged to get outside and get involved in their communities. To investigate how their coast is changing, for example, students visit their local beaches, take measurements of the height and slope of the sand and shore, and analyze their data using tools and methods similar to those used by coastal geologists.

Data collection

We developed an online data analysis tool called the Beach Profile Grapher, which allows students to input data about their local beach, compare their data with data collected in prior years, and look for evidence of change. Using this online tool in combination with the physical collection of data provides a powerful experience of monitoring and modeling places that are important to students.

At one study site on the northern coast of Kodiak Island, Alaska, students gathered data on White Sands Beach. A class of 25 eighth grade students used a GPS device and a georeferenced map of the beach, marked with four straight-line paths (called transects) highlighted along the beach (Figure 2). The transects matched those measured by scientists from the University of Alaska Anchorage (UAA) during the prior year (Figure 3).

Students used the Emery Method to create a profile of White Sands Beach, recording the change in elevation along each transect of the beach (Figure 4). Using this method, they placed one

Figure 3. Students used a georeferenced map to ensure they were in the same location as the University of Alaska Anchorage survey team.

Figure 4. Students collected data using the Emery Method on White Sands Beach.

rod (the landward rod) at the starting point of the transect and moved a second rod (the seaward rod) towards the ocean two meters away. The student holding the landward rod aligned the horizon and the top of the lower of the two rods, and recorded the change in elevation between the two (Figure 4). Next, students moved the landward rod to the position of the seaward rod, then moved the seaward rod toward the water and took new measurements. They repeated the process until the beach was profiled along each transect.

After student groups collected elevation change data at each of the four transects, they uploaded their data to the Beach Profile Grapher (Figure 5) in order to create, view, and compare line graphs of the elevation of a beach cross-section. The Beach Profile Grapher included pre-loaded data collected by the UAA survey team for five beaches in Alaska from 2018 to 2024. By overlaying graphs of the data the students collected with the UAA data, students were able to determine if White Sands Beach had been building or eroding along the same transect and if the beach was steeper or flatter along different transects.

Results

As a result of their work, the students saw that a cross-section of beach had been building up since 2019, yet in 2023 it had begun eroding. They discussed potential causes including recent storms, such as Typhoon Merbok, which hit Hooper Bay, another Alaskan classroom site, in September 2022. They also used computer simulations to model different scenarios and storm events (Figure 6). Students learned about coastal erosion mitigation strategies such as hard protection (concrete seawall) and soft protection (dune restoration), and asked their Elders and other community members what had worked in the past and what might work in the future. At the conclusion of the unit, students shared their data, alongside their observations, reflections, logbook notes, and stories of their beach with their community, including through song, dance, and video.

By providing meaningful activities that allow students to do the work of scientists and track how their coast is changing, the Precipitating Change unit helped students leave their own legacy. Future classes can use their data—and their stories—as resources and reference points to continue to study coastal change on their beaches.

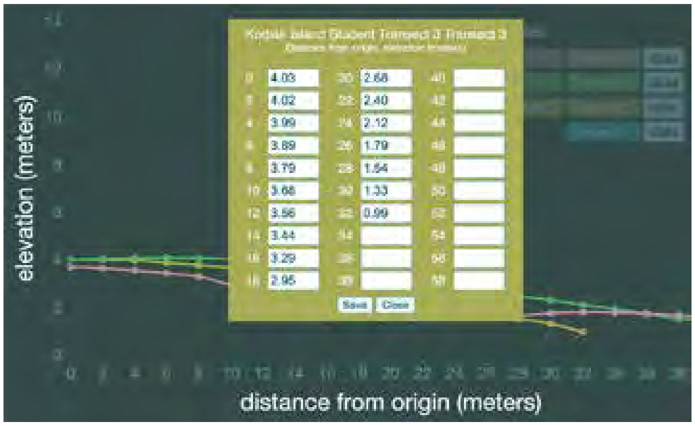


Figure 5. Student data for Transect 3 on White Sands Beach.

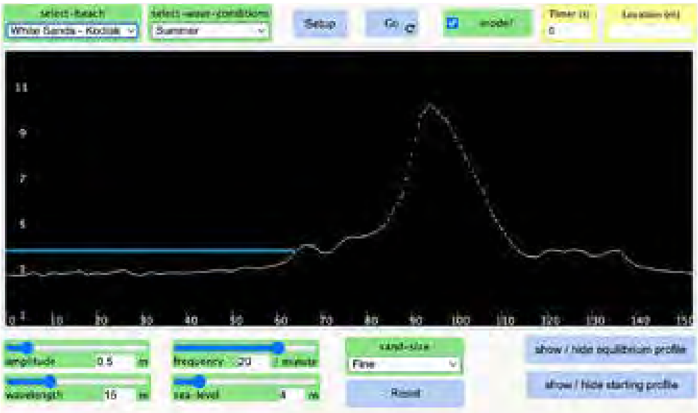


Figure 6. The Coastal Change Beach Model allows students to simulate different wave and sea level conditions that mimic storm surges in the area.

LINKS

Precipitating Change with Alaskan Schools—**concord.org/pc2ak**

Under the Hood:

Exploring Markov Chains with Animated Graphs in CODAP

Doug Martin
(dmartin@concord.org) is a principal software engineer

By Doug Martin

Suppose you're modeling the weather, and in your model the weather can be in three states: sunny, cloudy, or rainy. If it's sunny today, there are different probabilities that tomorrow it will remain sunny, become cloudy, or turn rainy. In this model, tomorrow's weather depends only on today's conditions, not on what the weather was like in the past. Implementing this simple three-state model using a "Markov chain" enables you to generate a large number of sequences of daily weather conditions. And those sequences can help you look for patterns, which is the beginning of machine learning, used in many artificial intelligence applications.

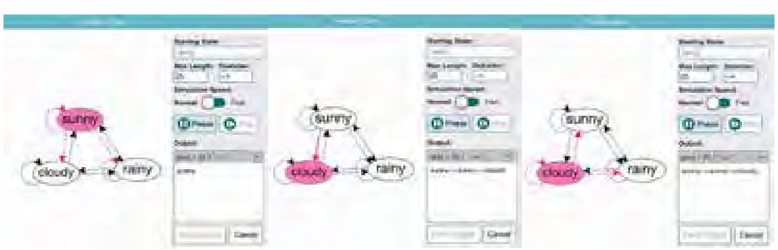
Markov chains model systems that move between different states, where the next state depends only on the current state and the preset transitional probabilities, and not on how the system arrived at that state. Our new Markov chain CODAP plugin—created for our AI Education Across the Curriculum project—allows students to generate Markov chain graphs from existing data or by dragging and dropping states and transitions to create their own graphs.

By showing how the states are connected and then animating the transitions between the states in real time, students can more intuitively see how the system moves from one state to another based on predefined probabilities. This visual representation makes abstract concepts such as state transitions and probabilistic outcomes more tangible.

Students can explore the graph by animating the sequence generation, either by playing the whole animation or stepping through it (Figure 1). At the start of the animation, the initial state is highlighted along with the transitions to the possible next states. As the animation progresses, each state transition is highlighted, providing a clear step-by-step illustration of how Markov chains function.

The animation algorithm starts by pre-generating the full sequence of states using the graph and starting state, as well as a maximum sequence length set by the student. Given a starting state, the transitions between states are determined by analyzing the graph and using the transition probabilities to determine the next state. The animation ends when either the sequence length has been reached or a state with no transitions outside of itself is reached.

Figure 1. Three states in the Markov chain animation. The dotted arrows represent the possible next state.



Once a sequence is generated, the algorithm then simply steps through that list of animation actions, re-rendering the graph at each step with the different parts of the graph highlighted depending on the action (Figure 2). By pre-generating the sequence and list of actions, the computational load during real-time visualization is eliminated and the code that manages the automatic or step-through animation is greatly simplified.

Along with highlighting the states and transitions during the animation, each state's label (e.g., sunny, cloudy, rainy) is output in a list within the plugin with an optional delimiter between labels, allowing students to connect the current animation state to the sequence output. At the end of the animation, the full sequence of states is displayed in a CODAP table, where students can analyze the data in detail, looking for patterns and performing statistical analyses on the state transitions.

Combining dynamic visualization with the data analysis capabilities of CODAP creates a highly effective tool for exploring Markov chains.

Figure 2. The generateSequence()function returns a path as a sequence of states (called nodes) through the Markov chain starting at either a selected starting state or a randomly chosen state.

```
const generateSequence = (graph: GraphData,
options: GenerateOptions) => {
  const {startingNode, lengthLimit} = options;
  const generatedResult: Node[] = [];

  let currentNode = startingNode ||
chooseRandomNode();
  while (currentNode && generatedResult.length
< lengthLimit) {
    generatedResult.push(currentNode);
    const currentEdge =
chooseRandomEdge(currentNode);
    if (currentEdge?.to) {
      currentNode = graph.nodes.find((iNode) =>
iNode.id === currentEdge.to);
    } else {
      currentNode = undefined;
    }
  }

  return generatedResult;
};
```

LINKS

AI Education Across the Curriculum—concord.org/storyq2
CODAP—codap.concord.org
Markov chain—models-resources.concord.org/markov-chain-plugin/

Teacher Innovator Interview: Christina Chin

Middle school science teacher, San Jose, California

As a child growing up in Silicon Valley, Christina Chin was no stranger to technology. Her father, an "old IBMer who brought home all kinds of gadgets," gave Christina her first laptop, a bulky briefcase-like device. She still remembers the screen's distinct monochromatic orange pixels. These days, she's using considerably more modern technology with her middle school students to get them excited about data science.

Christina did not start out in the classroom, nor did her original career path include computers. Trained in genetics, she began work in pharmaceutical research. She recalls spending her days "isolated in a lab with no windows." When she decided to pack up her lab coat, she moved to Hong Kong. "I wanted to find my roots," she laughs.

While overseas, she married and had two children. She also received a master's degree and teaching credentials through National University, taking online classes in the "infancy of distance learning" when there were assigned readings and discussion forums.

As a program requirement she completed her student teaching in the U.S., but her first teaching job was back in a local Hong Kong school. She also taught in Australian and Singaporean international schools, where she was impressed by the elementary science curriculum. "It spirals. It's age appropriate. It emphasizes precision with vocabulary."

Her family eventually moved back to California, where she is now in her fourteenth year teaching fifth through eighth grade. Her STEM school encourages teachers to try new things. When Christina learned about the Boosting Data Fluency project, a collaboration between WestEd and the Concord Consortium designed to increase teacher and student facility with exploring large datasets to learn science and math, she was eager to participate. "Clearly, data science is an up-and-coming thing," she explains.

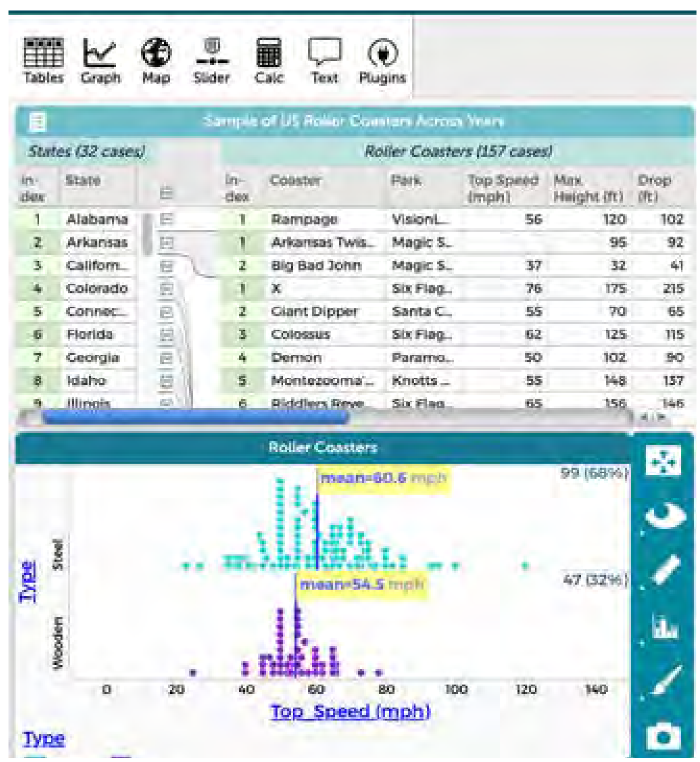
But she found that getting her students to explore real-world datasets required stepping stones, including critical support in math. Her fifth graders didn't know the difference between a line graph and a bar graph, for example. Christina now explicitly teaches vocabulary around patterns and trends. And she starts small, beginning with small datasets and having students graph the points by hand, then moving to Excel, where students practice graph interpretation skills, and finally to CODAP.

Last year, she gave her eighth grade students a large dataset about roller coasters. Because they were tech-savvy, she first had them complete the CODAP tutorial on their own before getting into the roller coaster data. She says her students really enjoyed the drag-and-drop features and were able to easily make graphs to compare the top speeds of wooden and steel roller coasters, for instance. "Their explorations were really organic," she says,

beaming with pride. CODAP, she adds, "really showed students the power of using the right tool to crunch numbers really fast and do something with it."

Christina equates data fluency with critical thinking and civic action. Her students look at data "all the time" in their science class, and it's important to her that they understand where the data comes from. By teaching students to understand data, she wants to empower them to stop and ask, "Wait, where is that data coming from? Wait, why are you making that decision? Is that the right decision for us?" Being able to ask questions about and with data is the key to data fluency. Christina is thrilled to boost her students' abilities.

Data about roller coasters and a graph of the mean top speed of wooden and steel roller coasters displayed in CODAP.



The Concord Consortium

25 Love Lane, Concord, MA 01742

The Concord Consortium is happy to announce the following new grants from the National Science Foundation.

The Pelehonuamea Project

We are expanding an existing Research-Practice Partnership (RPP) to include Indigenous Hawaiian students in order to broaden their sense of agency and investment in learning outcomes. The RPP will include middle school students in the co-design of a curriculum module focused on volcanic risks and hazards. To contextualize and situate learning about volcanic eruption from the Mauna Loa volcano, students will conduct ethnographic studies to capture Hawaiian historical and lived knowledge about volcanic eruptions. They will connect these narratives to the results of visual block-based coding to model lava eruptions and their impacts. Research focuses on 1) how experiencing a culturally and geographically relevant integrated geoscience and computer science curriculum module affects students' attitudes towards computer science and computational thinking, and 2) to what extent students build computer science and geoscience knowledge to better prepare for diverse job opportunities.

Data Science Interventions Connected to University Athletics

We are partnering with the University of Maryland, College Park to conduct groundbreaking research on how Black male athletes engage with data science and the systemic barriers and biases they face in STEM education. The project will co-design, implement, and assess the impact of a potentially transformative higher education intervention called DataGOAT (Greatest Of All Time) to enable Black male football and basketball athletes to study their own personal sports data and develop STEM identities through critical data literacy. By connecting sports performance and health to data analysis and visualization, we aim to facilitate career pathways from sports to STEM.

Open Data Exploration Tools for K-12 Education

A new initiative is taking the first step towards developing an open-source ecosystem designed to enhance data science education in K-12 schools. We aim to engage and expand the CODAP community, including code contributors, plugin developers, educators, curriculum developers, translators, researchers, and others; identify a distributed development infrastructure for supporting users; research organizational, governance, and coordination models; and explore avenues for ensuring a sustainable ecosystem for CODAP, including commercial opportunities such as professional development services, foundation support, or partnerships with educational technology or industry supporters.

Designing Interactive Visualizations of Neural Pathways

Artificial intelligence (AI) is transforming numerous industries and catalyzing scientific discoveries and engineering innovations. To enter an AI-ready workforce, young people must be introduced to core AI concepts and practices. Our exploratory project will develop and test an interactive, dynamic digital learning tool for middle school students to learn how to interpret neural networks and collaborate with the algorithm to improve AI systems. The team includes learning experience designers and technology developers from the Concord Consortium, computer scientists from Carnegie Mellon University, educational researchers from North Carolina State University, curriculum specialists and teacher educators from Mississippi State University Center for Cyber Education, and usability and feasibility evaluators from WestEd.

Data By Voice

Data science has become essential in modern society, with growing career opportunities and widespread adoption in educational curricula. However, blind and low-vision (BLV) students often lack the tools necessary for meaningful engagement with data. In collaboration with Perkins Access Consulting at the Perkins School for the Blind, we aim to enable BLV students to interact with data through voice commands, sonification, and AI-generated audible descriptions. Leveraging a cutting-edge large language model from generative AI technologies, we are developing and researching an AI-powered agent embedded in CODAP, which provides the interface between the user, the generative AI model, and CODAP. The plugin interprets BLV users' voice commands to perform data transformations, generates data representations, facilitates non-sequential navigation and exploration of data representations, and provides verbal and sonified descriptions of data representations.

Editor: **Chad Dorsey** | Managing Editor: **Cynthia McIntyre** | Design: **Jan Streitburger** | **concord.org**

The Concord Consortium | Concord, MA | El Cerrito, CA | 978-405-3200 | fax: 978-405-2076

@Concord is published two times a year by The Concord Consortium, a nonprofit research and development organization dedicated to innovating and inspiring equitable, large-scale improvements in STEM teaching and learning through technology.
This material is based upon work supported by the National Science Foundation under several grants (DRL-2148451, DRL-2200887, DRL-2241021, DRL-2048463, DRL-2101198, DRL-2241669, DRL-2101049, TI-2346126, DRL-2418281, DUE-2411706, DRL-2405615, and DRL-2434388).