

Description of Activities

1. Gotham Data Clinic

Gotham Data Clinic is a New York City-based charitable organization whose mission is to develop high-quality, openly licensed educational material in computing and data science, build technology to support its delivery, and train the next generation of scientists and technologists as well as the general public in these computational and data science skills.

The field of data science is a relatively new field and has largely been developed in the private sector. It extends the use of statistics and exploratory data analysis to understand and make inferences of the world. These skills are necessary across all fields of inquiry as the world is becoming ever-increasingly data-driven. Academic data science is burgeoning, growing from 13 programs in 2014 to over 50 in 2022 [1] but it is still not sufficient to meet the growing needs. Gotham Data Clinic aims to partner with colleges and universities as well as local community organizations to develop content and curriculum to meet the needs for data science training.

2. Founding and Mission

Gotham Data Clinic was born out of a National Institutes of Health Science Education Partnership Award granted to NYU to develop and deliver a neuroscience curriculum to high school students in NYC. We created a research platform to connect EEG hardware to students' computers to show them their brain activity for the first time, develop custom experiments, and run studies the same way researchers would in a laboratory setting. We also taught students how to analyze the results of their experiments using data science methods. This curriculum, BrainWaves [2], provided hands-on experiential learning to thirty different high schools across the city.

The grant period for the award was only for five years. We wanted to find a home for this research platform and its curriculum after the grant period ended so we established this nonprofit to be the stewards of the program and to be a place to incubate new programs to teach computational education. We have begun to extend our curriculum offerings to teach data science, machine learning, and artificial intelligence.

Our vision is to inform, prepare, and train the next generation of scientists and technologists and the broader public in computational and data sciences for a more fair and responsible future.

Our mission will be to

- Develop next-gen computational and data science education content and platform.
- Lead workshops and training on computing and data science.
- Publicize our work through scholarly engagement as well as public engagements.

3. Specific Activities

We think the best way to realize our vision of a fairer and more responsible future is to build new tools and curricula to educate, inform, and empower the leaders of tomorrow.

A. Developing Educational Materials

Gotham Data Clinic has delivered workshops on data literacy. We've partnered with local organizations to teach the public how to make sense of data, understand where it comes from, and make decisions from the information.

Gotham Data Clinic has been developing open-access curricula for computing and data science education. These courses teach the fundamentals in scientific computing, exploratory data analysis, statistical inferences, and machine learning techniques. We are partnering with local colleges and universities to deliver this material. This fits in with our overall strategy to prepare learners in computing and data science skills.

It is intended that 34 percent of the organization's time will be dedicated to research and development of educational content.

B. Developing Education Technology

We are actively working to build an education hub to house all of our educational efforts in one place and provide a computational education for all learners on our platform. Using some recent advances in the scientific computing ecosystem, there are

now tools to make setting up Python and R, two essential programming languages for data science, easy to set up and access, all in a web browser. This removes the need to cover server cost, which are typically associated with running these computational environments. Our hub will greatly increase access to data science education while dramatically reducing the maintenance costs.

We plan to provide a hosted solution for different communities and provide service support for their educational and technical needs. We will also provide a self-hosting option so communities can have full control over their data and content. We plan to have community spaces on our hub to host and share course content among instructors. This would create a Github-like environment to share, remix, and modify lessons and create a forum for instructors to collaborate with one another.

It is intended that 33 percent of the organization's time will be dedicated to technology development.

C. Scholarly and Community Engagement

Gotham Data Clinic is currently working on a landscape analysis and review of data science education. This scholarly research gives us a perspective of the latest implementation work happening in secondary education and higher education sectors. Our aim is to continue to practice in both scholarly research when necessary as well as develop tools and curricula.

We plan to make our findings available and publicize all our work through public engagements. This includes participating in panel discussions, presentations, and invited talks.

It is intended that 33 percent of the organization's time will be dedicated to scholarly and community engagement.

[1] <https://magazine.howard.edu/stories/beyond-the-numbers>

[2] <https://wp.nyu.edu/brainwaves/>