Teaching Data Science and Building Data Literacy – Nationally and Globally

With the Northeast Big Data Innovation Hub & National Student Data Corps

Florence D. Hudson, Executive Director, PI, Founder of NSDC
Emily Rothenberg, NSDC Program Manager
National Student Data Corps
Northeast Big Data Innovation Hub at Columbia University
Northeast Big Data Innovation Hub Team

Florence D. Hudson
Executive Director
Northeast Big Data Innovation Hub

Lauren Close
Operations and Communications Manager

Emily Rothenberg
National Student Data Corps Program Manager

Sanket Bhandari
MS Data Science ‘23

Sneha Dahiya
MS Business Analytics ‘23

Anna Eggers
BA Film & Writing ‘24

Stephanie Guo
MS Business Analytics ‘23

Hannah Yang
MS Data Science ‘23

Varalika Mahajan
MS Data Science ‘23

Lylybell Teran
MS Data Science ‘23

Columbia University Students - Part Time
The Northeast Hub is a community convener, collaboration hub, and catalyst for data science innovation in the Northeast Region and around the world.

Building a diverse, equitable, and inclusive community with accessible resources is central to our work.

We are a community of over 8,300 individuals from 1,375 organizations, all 50 U.S. States and 61 countries. Join us!

Four Focus Areas guide our community engagement and activities.

- Education + Data Literacy
- Health
- Urban to Rural Communities
- Responsible Data Science: Security + Privacy + Ethics

Funded by NSF grants #1550284, 1748395, 1916585, 2028999, 2139391, and NIH AIM-AHEAD.

http://nebigdatahub.org/about/
Northeast Hub Focus Areas Guide our Activities

The NSDC is the Hub’s premier Education & Data Literacy Program

Education & Data Literacy
- National Student Data Corps (NSDC)
- Open/Share - Call for Participation

Data Science Resource Repository (DSRR)
- NSDC Global Chapter Community
- Data Science Panel Video Series
- Data Science Mentorship Program
- Education & Data Literacy Seed Fund Projects

Health
- COVID Information Commons (CID)
- CIC Student Paper Challenge
- CIC Working Groups
- Clinical IoT and TIPPS at HIMSS
- Blockchain & Data Science in Healthcare - Call for Papers
- Health Seed Fund Projects

Urban to Rural Communities
- Smart Cities Data Exchange
- Driver Video Privacy Challenge
- Transportation Projects
- Smart Cities and Buildings
- Urban to Rural Communities Working Group
- Urban to Rural Seed Fund Projects

Responsible Data Science
- Cybersecurity Risk Initiative
- Responsible Data Science Comic Books
- IEEE/OSA P2933 TIPPS Standard Working Group
- IEEE Connected Healthcare Cybersecurity Workshops
- Framework for Interpreting Data Equity Systems (FIDES)
- Data Sharing and Cyberinfrastructure WG
- Responsible Data Science Seed Fund Projects

http://nebigdatahub.org/about/
The National Student Data Corps (NSDC) is a community-developed initiative that provides resources and opportunities for students to learn data science, in a community of support, with a special focus on engaging underserved institutions, students, and communities.

There are three teams from the NSDC Founding Committee leading the effort.

- Content and Pedagogy Team
- Peer Instructor and Mentoring Team
- Outreach Team

https://nebbigdatahub.org/nsdc
As of June 2023, the NSDC community is made up of 4,430 individuals from 613 institutions representing the U.S. and 28 other countries – Australia, Bangladesh, Botswana, Canada, China, Ecuador, Ethiopia, France, Germany, Ghana, Hungary, India, Israel, Japan, Kenya, Lebanon, Mali, Mexico, Moldova, Nepal, Netherlands, Nigeria, Peru, Philippines, South Korea, Turkey, United Arab Emirates, and the United Kingdom.

https://nebigdatahub.org/nsdc
NSDC Provides over 200 Free Online Data Science Resources for Learners & Educators

https://nebigdatahub.org/nsdc
Data, Responsibly Comic Book Series

Teaching data science ethics and fairness. In English, Spanish, French, Portuguese, and Greek.

Included in the NSDC Educator and Learner Central, and Data Science Resource Repository (DSRR) (https://nebigdatahub.org/learner-central-data-science-ethics/)

Start or Join an NSDC Chapter! In US and Globally!

Chapter members benefit from:
- Data science resources
- Mentorship
- Research, volunteer and career opportunities

https://nebigdatahub.org/nsdc/nsdc-chapters/
NSDC Data Science Career Panels -
14 panels to date, 40 expert panelists, 3800 viewers!

Watch past panels on demand here: bit.ly/nsdc-panels
NEW NSDC MasterClass Series

The new NSDC MasterClass Series is designed to showcase how data science is being leveraged in various domains.

Highlighted domains may include:

- Health
- Finance
- Athletics
- Entertainment, Music & Media
- Technology
- Education
- And more!

https://www.youtube.com/@NortheastBigDataInnovationHub
The NSDC DSP provide **practical data science project experience** for learners of all skill-levels and education-levels, and highlight data science as a team effort!

Through a guided **Google Colab Notebook**, and supportive weekly mentorship calls, cross-institutional teams of students learn data science techniques and gain leadership and professional development skills.

**NSDC Data Science Projects (DSP)**

Join an upcoming NSDC Data Science Symposium!

The NSDC Data Science Symposium (DSS) invites students to learn about data science research and grow their research community.

Students from higher ed institutions are invited to create a data science research project and showcase their findings in a virtual poster board presentation.

The NSDC Data Science Symposium is held annually, and we welcome you to join the upcoming challenges!

https://nebigdatahub.org/nsdc/2023-nsdc-dss/
NSDC Mentorship Program

Find or become an NSDC Mentee or Mentor to support a data science journey! The National Student Data Corps Data Science Mentorship Program facilitates mentorship opportunities through the NSDC Slack Channel for over 610 data science students, professionals, and enthusiasts across the U.S. and beyond!

Interested in becoming an NSDC Mentor?
bit.ly/become-mentors

Interested in becoming an NSDC Mentee?
bit.ly/become-mentees

https://nebigdatahub.org/nsdc/nsdc-mentorship-program/
Data Science Resource Repository

The DSRR is an open access platform with 800+ resources including data science best practices and content to broadly increase knowledge capacity, with a specific focus on underserved communities.

Topical areas include:

- Accessibility & Data Science
- AI
- Apache
- Azure
- Big Data
- Bioinformatics
- Cities
- Climate Change & Environment
- Coding
- COVID
- CSS
- Cybersecurity & Privacy
- Data Analytics
- Data Modeling
- Data Science Basics
- Data Science Careers
- Data Science for Good
- Data Visualization
- Database Management
- Deep Learning
- Diversity
- Ethics & Bias
- Java
- Kubernetes
- Linear Algebra
- Machine Learning
- Math Fundamentals
- Natural Language Processing
- Pandas
- Power BI
- Probability
- Public Health
- Public Policy
- Python
- Quantitative Reasoning
- Quantum Computing
- Research Tools
- Rural Populations
- Splunk
- SQL
- Statistics
- Tableau
- TypeScript
- Water

Volunteer opportunity to add more content, including Spanish content.

https://nebigdatahub.org/data-science-resource-repository/
**Data Science Resource Repository**

Resources are grouped by subjects, or topical areas.

Users can browse, search, and filter resources by keyword or tag.

Learning formats include articles, journals, tutorials, videos, textbooks, courses, challenges, and podcasts.

OpenDS4All - Call for Participation

The NEBDHub & NSDC, in collaboration with the Linux Foundation AI & Data, is delighted to invite you to join the leadership and development team of the OpenDS4All GitHub repository.

The OpenDS4All GitHub repository includes hundreds of resources from slide presentations on data science ethics and supervised machine learning, to Jupyter notebooks which provide hands-on learning opportunities.

If you would like to join this volunteer team, the three volunteer roles include:

- Technical Steering Committee Member
- Content Review Team
- Content Developers

https://nebigdatahub.org/opends4all/
Thank you!

Florence D. Hudson, Executive Director, PI, Founder of NSDC; fh2417@columbia.edu
Emily Rothenberg, NSDC Program Manager; er3101@columbia.edu
National Student Data Corps
Northeast Big Data Innovation Hub at Columbia University

- Sign up for the NEBDHub & NSDC mailing list: bit.ly/HubNewsSignup
- Join the NSDC Community: nebigdatahub.org/nsdc
- Join the NSDC Slack Channel: bit.ly/nsdc-slack-channel
- Keep up with the NEBDHub & NSDC on the NEBDHub YT Channel: youtube.com/@NortheastBigDataInnovationHub
- Join a Hub or NSDC event: nebigdatahub.org/events-calendar
- Any questions, reach out to us at contact@nebigdatahub.org

All available through: https://nebigdatahub.org/about