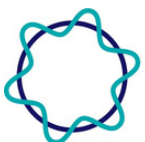# Academic Data Science Alliance

Founding Member Book
2021-2022

Academic
Data Science
Alliance

April 27, 2022

2021-2022 ADSA Member Book

The Academic Data Science Alliance is a community network for data science leaders, practitioners, and educators who take responsibility for a just, equitable future where data science approaches are thoughtfully applied in all domains for the benefit of all.

The ADSA Member Book highlights the institutions who: have provided funds for ADSA through membership dues; support ADSA as a leading organization for academic data science; and agree with the guiding principles in ADSA's Mission, Vision, Values. ADSA member institutions represent a range of models, maturity levels, and target audiences at academic institutions in the US and beyond. **This inaugural Member Book showcases our Founding Members**. In the years to come, we will continue to offer institutional membership and Member Books to provide a resource for students, administrators, and leaders of data science programs.

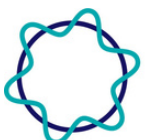ADSA is generously supported by the following organizations:

Academic Data Science Alliance

www.academicdatascience.org

2021-2022 ADSA Member Book

Table of Contents

www.academicdatascience.org

2021-2022 ADSA Member Book

Table of Contents (continued)

2021-2022 ADSA Member Book

Table of Contents (continued)

2021-2022 ADSA Member Book

Table of Contents (continued)

www.academicdatascience.org

American Statistical Association

SUMMARY

The American Statistical Association is the world's largest community of statisticians and data scientists. It is the second-oldest, continuously operating professional association in the country. Since it was founded in Boston in 1839, the ASA has supported excellence in the development, application, and dissemination of statistical science through meetings, publications, membership services, education, accreditation, and advocacy.

SOCIAL

email: asainfo@amstat.org
web: www.amstat.org
twitter: @AmstatNews
linkedin:
www.linkedin.com/company/american-statistical-association---asa
facebook:
www.facebook.com/AmstatNews
instagram:
www.instagram.com/amstatnews
pinterest:
www.pinterest.com/amstatnews

LOCATION

American Statistical Association
732 North Washington Street
Alexandria, VA 22314-1943

PROGRAM OVERVIEW

ASA is a professional society. It does not grant degrees. Rather it supports departments, programs and individual professionals through journals, meetings, professional development, networking and advocacy.

ASA champions the data science community in a variety of ways beyond its membership in ADSA. For instance, ASA is a member of CSAB, which as part of ABET accredits undergraduate programs in data science. ASA provides input into the accreditation criteria and ASA members are trained to serve as program evaluators.

ASA has sections which provide members with the opportunity to network, research, and develop as professionals with others who are interested in the same topic. Our largest section is Statistics Learning and Data Science. Our Statistics and Data Science Education section is also large and active.

ASA publishes or co-publishes 16 leading journals in statistics and data science, Statistical Analysis and Data Mining, the Journal of Computational and Graphical Statistics, and the Journal of Statistics and Data Science Education.

ASA's professional meetings are highly regarded. The flagship meeting for the statistics profession is the Joint Statistical Meetings, held in late July or early August. The ASA manages this meeting, which includes representation from 10 other statistical and data science organizations. The Symposium on Data Science and Statistics (held in May) and the Women in Statistics and Data Science Conference (October) are among our other high-quality meetings.

ASA promotes data science and statistics in the statistical agencies of the federal government and in state and local government as well through advocacy and through our "Count on Statistics" platform.

ASA supports students through competitions and outreach. ASA DataFest is an annual 48-hour competition in which teams of undergraduate students work to reveal insights into a rich and complex data set. ASA hosts data challenges in the fall and spring through our "*This* is Statistics" program, which is our outreach to high school and undergraduate students.

ASA also supports K-12 and college instruction through providing guidelines for teaching at various levels, resources for teaching, and networking for instructors.

ASA provides professional development through in-person and online instruction in a wide variety of topics. Professional development coursework spans everything from basic overviews to advanced instruction in cutting edge topics, and includes courses in non-technical skills as well.

The board and staff leadership of the ASA welcomes suggestions for additional ways to support and advocate for the data science community.

Academic
Data Science
Alliance

Data Science @ AUP

THE AMERICAN UNIVERSITY OF PARIS (AUP)

SUMMARY

Chartered as a liberal arts college in 1962, AUP is today an urban, independent, international university. Data science at AUP spans topics such as machine learning, artificial intelligence, algorithms, statistics and geographical information systems. Teaching and research cover a wide variety of application fields such as law, astrophysics, and literature.

PROGRAM STATISTICS

- Number of faculty: 4
- Affiliated faculty: 36

DEGREES, PROGRAMS, SPECIALIZATIONS

- MSc in Human Rights and Data Science
- Undergraduate minor in Data Science

SOCIAL

email: croda@aup.edu
web: www.aup.edu

LOCATION

5, Boulevard de La Tour-Maubourg
75007 Paris, France

DATA SCIENCE AT AUP

All AUP students are welcome and encouraged to engage with data science. These engagements may take different forms, from a master program in Human Rights and Data Science to an undergraduate minor in data science, from a course in Environmental Data Science to a student project on the use of paper at AUP, from an internship helping the CIO with data protection to the development of a Python notebook on natural language processing for

comparative literature students. Undergraduates at AUP may also take a data science introductory course to satisfy components of our Global Liberal Arts Core Curriculum.

MINOR IN DATA SCIENCE

Data science can be applied to all fields of human endeavor – such as understanding environmental data or clinical trial data, building economic models, undertaking psychological studies or political and marketing analyses, or simply making sense of the huge swathes of data generated by our social media. Everyone is familiar with examples of how the interpretation of this data may be affected by biases and errors that result in unexpected and undesirable effects on particular groups. AUP's Minor in Data Science gives students the opportunity to learn about this burgeoning area, taking a critical approach. Students in the minor learn how to apply the data science process to their field of interest, discovering along the way the Python programming language, the essential notions of statistics and probability which underlie data science, and the ethical concerns surrounding the application of data science.

Data Science @ AUP

THE AMERICAN UNIVERSITY OF PARIS (AUP)

MASTER IN HUMAN RIGHTS AND DATA SCIENCE (HRDS)

The MSc in Human Rights and Data Science at AUP prepares students to confront ethical questions at the forefront of emerging high-tech industries such as data protection and ethical AI, exploring how technological development may complement human rights protections rather than impinge upon them. The HRDS program combines a rigorous foundation in data science techniques with the legal and philosophical considerations necessary to ensure the ethical implementation of a data science project and the policies regulating it. The program concerns itself as much with the use of data centered software as it does with its development, with the human and social costs of emerging digital technologies as much as with the benefits.

A student project studying how food insecurity in Michigan is addressed by Commodity Crop Subsidies vs. Small Local Farms.

THE HRDS CURRICULUM

The curriculum includes classic courses in both data science and human rights law, as well as a series of innovative workshops, designed in collaboration with our partner institutions, giving students the opportunity to build a strong professional network and apply their knowledge on practical cases. Past workshops and information sessions have been run by lead data scientists, chief privacy counsels and data protection officers of large organizations, including the French Development Agency, MasterCard, International Committee of the Red Cross, OECD Health Statistics, French Data Protection Authority, among others. In addition to their academic diploma, students are guided through the preparation of a certification exam in data protection.

Thanks to their deep understanding of the fast-changing landscape of data-related technology and regulations and by embracing both the intellectual and vocational nature of human rights law and data science, AUP's HRDS graduates will help alleviate the considerable shortage of experts in ethical AI confronting government and industry.

Student projects exploring: (1) Ideal US states for cannabis cultivation and (2) Production and import of energy in US states

Belmont Data Collaborative

SUMMARY

A university wide initiative that looks to infuse data skills into every facet of the culture and curriculum within each major and college as well as creating opportunities for real world projects that give students and faculty experience as well as help the social well-being of the community.

PROGRAM STATISTICS

- Number of faculty: 3
- Number of staff: 3
- Affiliated Faculty: 10
- Students: 70 - BSA-BBA: 25 DS-BS, 25 - BSA-Minor, 20 - DS-Minor

DEGREES, PROGRAMS, SPECIALIZATIONS

- Data Science - Bachelor of Science
- Data Science - Minor
- Business Systems and Analytics - Bachelor of Business Administration
- Business Systems and Analytics - Minor

SOCIAL

email: data@belmont.edu
web: www.belmont.edu/data
twitter: @belmontdata
linkedin:
www.linkedin.com/company/belmontdata

LOCATION

1900 Belmont Boulevard
Nashville, TN 37212

PROGRAM OVERVIEW

The Belmont Data Collaborative works with all of campus to make sure

that every student at Belmont not only has a level of data literacy to fluency, but also each student become a data storyteller. This is done through not only curriculum, but events (data hackathons), projects that look to solve complex community and societal problems, and experiences through internships and study abroad.

INDUSTRY PARTNERSHIPS

To keep ahead of the trends, the Belmont Data Collaborative continues to bridge the gap between academia and industry through partnerships that look to include all types of industry partners (music, healthcare, finance, etc.). Therefore, through the BDc, industry professionals are deeply

ingrained in the curriculum and experiences with our students. Such as the partnership with the Mechanical Licensing Collective which provided data, expertise, and tours of their facilities to better equip our future music/data professionals.

CHAMPIONS FOR CHANGE

Becoming radical champions for change and bettering local and national communities is a key aspect to the Belmont Data Collaborative. Through a joint effort with several key leaders and groups within Nashville, TN, the BDC is creating a community level data warehouse that enables a data-driven approach to solving some

of societies hardest problems. The first one will focus on hypertension within distressed communities in Nashville. By taking a community level and data-driven approach, it allows for a social and well-being approach to service based solutions.

Academic Data Science Alliance

BU Faculty of Computing & Data Sciences

SUMMARY

Founded in 2019, the Faculty of Computing & Data Sciences (CDS) is a transdisciplinary, degree-granting academic unit that augments existing programs who remain under the traditional university organizational structure. We are propelling data sciences into the future and are actively recruiting faculty, admitting students, and moving to a new state-of-the-art building.

PROGRAM STATISTICS

- Number of CDS faculty: 5 core; 36 secondary
- Number of CDS staff: 10
- Number of affiliated faculty: 350+ Hariri Institute affiliates
- Number of students taught: 200+ per semester
- Number of Data Science concentrators: 40 and growing

DEGREES, PROGRAMS, SPECIALIZATIONS

- PhD in Computing & Data Sciences
- BS in Data Science
- Undergraduate Minor in Data Science
- All programs feature methodological and in-the-field tracks

LOCATION

665 Commonwealth Ave.
Boston MA 02215

PROGRAM OVERVIEW

Boston University (BU) is one of the largest private institutions in the US with 34,000+ students, 10,000+ faculty and staff, spread across 17 schools and colleges covering everything from Fine Arts to Medicine. The mission of CDS is to be an integrative cross-cutting academic unit, facilitating the movement of ideas, research, and educational programs across the University and building bridges between established departments, schools, and colleges.

THE BU CDS VISION

Our aim is nothing less than to democratize access to computing and data sciences. We do this by preparing a more diverse cohort of learners for rewarding computing careers, by catalyzing faculty research in disciplines that are yet to be transformed by the advent of machine learning and AI, and by bringing data-driven innovation to bear on societally relevant challenges that are often overlooked by industry. Through experiential learning programs, powered by BU Spark! (a technology incubator and experiential learning lab focused on interdisciplinary computing and data science projects), we help students realize their potential by providing access to resources, knowledge, and expert networks to support their innovation and applied learning journeys.
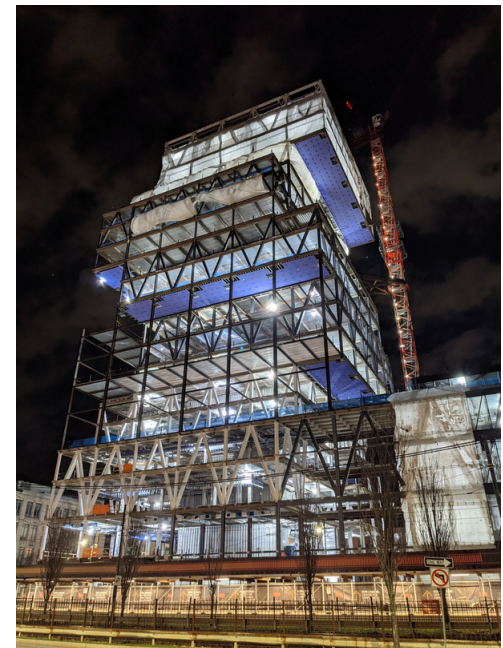
A DISTINCT ORGANIZATION

The model of CDS is nationally unique. Distinct from the existing cognate academic units at BU, CDS augments and complements these departments without transplanting them under a new school or college. This organization recognizes the importance (and preserves the diversity) of the established academic and professional cultures. Today, faculty from across the many schools and colleges at BU participate in this new dynamic initiative - providing access to resources, benefits, and a broad community of data scientists.

Faculty of Computing & Data Sciences

The CDS community is comprised of both faculty from cognate units and tenured and tenure-track faculty appointed through CDS with shared interests in the many aspects of data science. As such, CDS is a self-governing autonomous academic unit and not merely a division or institute that acts as an administrative overlay to coordinate the priorities of other academic units. Providing an academic home that is distinct from those supporting the core disciplines that comprise data science allows CDS to recruit truly interdisciplinary faculty who identify more with how data science scholarship is pursued as opposed to what scholarship is leveraged by data science.

CONNECTING TISSUE

We are the home for data science at BU, a deeply inclusive, collaborative and cooperative place with porous boundaries. A hub for data science, CDS is the convener and amplifier of university-wide interdisciplinary research centers and labs, such as the Rafik B. Hariri Institute for Computing, an incubator that accelerates data-driven research in various disciplines by leveraging BU's strengths in existing centers and new initiatives focused on data science, artificial intelligence, software engineering, computational science, cloud computing, digital health, data privacy, cybersecurity, and the nexus of computing, society, and law.

HUBS FOR IMPACT

Rather than clustering faculty, students, and programs around computing and data subspecialties (eg. machine learning, data mining, cloud computing), we center on thematic areas of impact, currently around Equity, Sustainability, Health, and Civic Tech. Much of our work is done through co-Labs that involve faculty, students, and external partnerships around applied research, curricular, and co-curricular collaborations. The experiential learning component of these collaborations is delivered through the BU Spark! program leveraging a shared infrastructure and staff support.

SOCIAL
email: cds-admin@bu.edu
web: bu.edu/cds
twitter: @bu_cds
linkedin:
www.linkedin.com/company/bucds
facebook: www.facebook.com/BU.CDS
instagram: @bu_cds

The Data Science Collaboratory
At Colgate University

SUMMARY

The collaboratory is a data analysis and collaboration network focused exclusively on smaller colleges and universities in New York State.

PROGAM STATISTICS

Number of core faculty: 2
Number of affiliates: 3

SOCIAL

email: datascience@colgate.edu
web: shiny.colgate.edu
twitter: @DataSciCollab

LOCATION

Case-Geyer Library
13 Oak Drive
Hamilton, NY 13346

PROGRAM OVERVIEW

The Colgate University Data Science Collaboratory was founded in 2018 by Joshua Finnell, Head of Research and Instruction in the University Libraries, and Will Cipolli, Assistant Professor of Mathematics with funding from the Central New York Library Resources Council New Initiatives Grant.

Our goal is to develop a vibrant research community among faculty and students. We aim to provide statistical guidance and resources to researchers across fields increasingly reliant on data science. Further, by incorporating students into the scientific community, they gain exposure to data science research and experience selecting, applying, and interpreting the results of appropriate techniques in various real-world contexts.

COLLABORATORY GOALS

The Data Science Collaboratory at Colgate University is focused on three interconnected goals:

- Introduce emerging student scholars to data analysis across multiple disciplines.
- Cultivate a community of data science researchers and scholars across the Central New York region for mentorship and collaborative projects.
- Develop statistical tools that lower the barrier to and increase the quality of quantitative research.

Academic Data Science Alliance

Data Science Institute

SUMMARY

DSI advances the state-of-the-art in data science; transforms all fields, professions, and sectors through the application of data science; and ensures the responsible use of data to benefit society. We train data scientists, develop innovative technology, foster collaborations, and work with industry to bring promising ideas to market.

PROGRAM STATISTICS

- Number of core staff:
  - 5 Research Scientists and Scholars
  - 18 Postdoctoral Researchers
- Number of affiliated members: 360+
- Number of students: 417 Current Graduate Students (Fall 2021)

DEGREES, PROGRAMS, SPECIALIZATIONS

- M.S. in Data Science
- Ph.D. with a specialization in Data Science
- Certification of Professional Achievement in Data Sciences
- Specialization in Executive Education

LOCATION

Northwest Corner Building
550 West 120th Street
Suite 1401
New York, N.Y. 10027

PROGRAM OVERVIEW

Drawing on Columbia's strengths in computer science, statistics, and industrial engineering and operations research, DSI was launched in 2012 to unite our expertise and a university-wide interest in this revolutionary approach. The university is a trailblazer in the field and is uniquely poised to expand data science to every corner of the institution.

We use the tagline "Data for Good" to capture succinctly the who, what, when, why, and how of data science at Columbia. We aim to have a positive impact on society by tackling societal grand challenges, such as climate change, health care, and social justice. Tackling such challenges cannot be done by one discipline alone, and given the kinds and amounts of data amassed in these sectors, data science will be at the heart of addressing these challenges.

DEGREE, CERTIFICATE, AND SPECIALIZATION DETAILS

M.S. in Data Science - This program is jointly offered in collaboration with the Graduate School of Arts and Sciences' Department of Statistics, and The Fu Foundation School of Engineering and Applied Science's Department of Computer Science and Department of Industrial Engineering and Operations Research.

Ph.D. with a specialization in Data Science - The Ph.D. specialization in data science is an option within the Applied Mathematics, Computer Science, Electrical Engineering, Industrial Engineering and Operations Research, and Statistics departments.

Certification of Professional Achievement in Data Sciences - This program is jointly offered in collaboration with the Graduate School of Arts and Sciences and The Fu Foundation School of Engineering and Applied Sciences. Join us from anywhere in the world as the program is now also offered online.

Executive Education - We partner with the Center for Technology Management to provide industry leaders with a better understanding of how to design and manage data science applications in their organizations.

SOCIAL

email: datascience@columbia.edu
web: datascience.columbia.edu
twitter: @DataSciColumbia
linkedin: www.linkedin.com/company/datascicolumbia
facebook: www.facebook.com/DataSciColumbia
instagram: www.instagram.com/datascicolumbia

Data Science Institute

COLUMBIA UNIVERSITY

THE JPMORGAN COLUMBIA UNIVERSITY DATA SCIENCE INSTITUTE MASTER'S DIVERSITY FELLOWSHIP AWARD PROGRAM

DSI offers fellowships for M.S. in Data Science students from historically underserved or underprivileged populations. Support is guaranteed for two semesters of full-time study, including full tuition, fees, and a stipend to cover living expenses in New York City. Fellows will also be invited to complete an internship with JPMorgan. Renewal for the third and final semester of the program is contingent on completing the internship. Please note: Fellowship applicants should be U.S. citizens or permanent residents.

Step 1: Review the M.S. in Data Science Eligibility and Application Requirements

- Undergraduate degree
- Prior quantitative coursework (calculus, linear algebra, etc.)
- Prior introductory computer programming coursework
- Online application
- Personal statement
- Uploaded transcripts from every post-secondary institution attended
- Three recommendation letters
- Curriculum vitae / résumé
- Official Graduate Record Examination (GRE) General Test Scores are optional for Spring and Fall 2021 applications
- $85 non-refundable application fee
- TOEFL, IELTS, or PTE Academic test scores, if applicable

Step 2: Apply to the M.S. in Data Science Program through Columbia's Fu Foundation School of Engineering and Applied Science.

Step 3: Apply for the JPMorgan Fellowship
Prospective students must submit a 250 – 500 word reflection that provides insight into how their identity, life experiences, and perspective would advance diversity in the data science community. The prompt for the reflection is available within the application platform.

WHAT PEOPLE ARE SAYING ABOUT DSI

"Our Data Science Institute is indispensable to virtually every scholarly initiative at the University dedicated to addressing a societal problem."
- Lee C. Bollinger, Columbia University President

"I love how students are encouraged to branch out across departments and engage in projects with faculty and other students across the university."
- Kailande Cassamajor, Class of 2023, B.S., Biology and Psychology, Howard University

"Columbia offers you everything you need. If you have a love for learning, it's the perfect soil for you."
- Jack Yang, Class of 2022, Managing Director @ JPMorgan

"I found it to be a well rounded program—not just technical, but also applied. I appreciate how students are given the opportunity to use cutting edge tools on real problems."
- Alberto Munguia Cisneros, Class of 2021, VP, Risk Management @ Morgan Stanley

"I ended up picking Columbia for the rigorous academic curriculum and the fact that the New York City area is a prominent tech hub."
- Carlo Provinciali, Class of 2020, Data Scientist @ Latch

"DSI honed my skills in not just the theory and application of data science concepts, but gave me multiple opportunities to contribute to my community and step up to take on more responsibility beyond my academics."
- Aakanksha Joshi, Class of 2019, Data Scientist @ IBM

Academic Data Science Alliance

SUMMARY

The Institute for Data Engineering and Science supports research in data science foundations and data-driven discovery. Foundational areas of focus include machine learning, artificial intelligence, high-performance computing, algorithms, statistics, and optimization. The institute supports data-driven research in many areas including astrophysics, chemistry, biology, medicine, materials science, energy, and smart cities.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Master of Science in Analytics (MS Analytics)
- Master of Science in Analytics - Online (OMS Analytics)
- Doctor of Philosophy in Machine Learning (ML Ph.D.)
- Computational Data Analysis Minor
- Computing and Intelligence Minor
- Bachelor of Science in Industrial Engineering - Analytics and Data Science

PROGRAM STATISTICS

- Number of Faculty: 50
- Number of Staff: 12
- Affiliated Faculty: 200
- Number of Students:
  - M.S. Analytics - on-campus (154), online (5,239); Ph.D. in ML (123)

LOCATION

Georgia Institution of Technology
CODA Building, 12th Floor
756 W Peachtree St NW
Atlanta, GA 30308

PROGRAM OVERVIEW

The Institute for Data Engineering and Science (IDEaS) is one of Georgia Tech's ten Interdisciplinary Research Institutes, providing the thought leadership, expertise, and coordination for strengthening big data research across campus. The institute faculty and its activities span all colleges and disciplines at Georgia Tech. Constituent and affiliated centers of the institute include the center for High Performance Computing (CHiPC), the center for Research in Novel Computing Hierarchies (CRNCH), the center for Machine Learning (ML), Algorithms and Randomness Center (ARC), the industry funded center on Machine Learning for Seismic Industry (ML4Seismic), and the industry funded center for cloud computing (CloudHub). IDEaS also operates major NSF funded centers including the NSF South Big Data Regional Innovation Hub (SouthBDHub), the NSF Transdisciplinary Institute for Advancing Data Science (TRIAD), and the NSF AI Institute for Adult Learning and Online Education (ALOE).

Academic programs related to data science at Georgia Tech are managed by academic units and interdisciplinary programs. Chief among them relevant to data science include the Ph.D. degree programs in Machine Learning (ML) and Computational Science and Engineering (CSE), the M.S. degree programs in Analytics (on-campus and on-line), Urban Analytics, and Computational Science and Engineering (CSE). Tech also offers undergraduate minors and specializations in Computational Data Analysis (CS), Computing and Intelligence (CS), and Analytics and Data Science (ISyE).

SOUTHBDHUB

IDEaS operates the NSF South Big Data Regional Innovation Hub, in collaboration with the Renaissance Computing Institute. The Hub serves the Southern U.S. census region (16 Southern states from Delaware to Texas, Washington DC, Puerto Rico, the U.S. Virgin Islands, and territories). The Hub nurtures multi-stakeholder big data partnership to find data-driven solutions to regional, national, and societal challenges. Areas of particular focus include health disparities, environment, materials and manufacturing, smart cities and communities, team science, data sharing and cyberinfrastructure, and data science education and workforce development.

Academic Data Science Alliance

Georgia Tech
Institute for Data Engineering and Science

The Hub has an extensive network of more than 1300 members from universities, government labs, corporations, and foundations. For more information, please visit https://southbigdatahub.org/.

MACHINE LEARNING PH.D. PROGRAM

The machine learning (ML) Ph.D. program is a collaborative venture between Georgia Tech's colleges of Computing, Engineering, and Sciences. The curricular requirements include four core and five elective courses, a qualifying exam, and a doctoral dissertation defense. Students are admitted through one of eight participating home units. More information can be found at: https://ml.gatech.edu/phd.

ANALYTICS M.S. PROGRAM

The Master of Science in Analytics is an interdisciplinary analytics and data science program that leverages the strengths of Georgia Tech in statistics, operations research, computing, and business by combining the world-class expertise of the Scheller College of Business, the College of Computing, and the College of Engineering. By blending the strengths of these nationally ranked programs, graduates will learn to integrate skills in a unique and interdisciplinary way that yields deep insights into analytics problems. Students can join the program on-campus, or through the online offering. More information can be found at: https://www.analytics.gatech.edu/.

COMPUTATIONAL SCIENCE AND ENGINEERING M.S. AND PH.D. PROGRAMS

The Computational Science and Engineering (CSE) M.S. and Ph.D. programs are interdisciplinary programs offered by the College of Computing, the College of Engineering, and the College of Sciences. These graduate programs emphasize the integration and application of principles from mathematics, science, engineering and computing to create computational and data-driven models for solving important real-world problems. Students will be required to obtain a breadth of knowledge across a set of core areas, depth of knowledge in a specific computational specialization, and knowledge to apply computational and data-driven techniques in a domain of application. For the Ph.D. program, students will be expected to create significant computational artifacts (e.g., software), and complete independent research that advances the state-of-the-art. More information can be found at: https://cse.gatech.edu/academics.

URBAN ANALYTICS M.S. PROGRAM

The M.S. in Urban Analytics (MSUA) program is a joint venture between the School of City and Regional Planning, the School of Industrial and Systems Engineering, the School of Computational Science and Engineering, and the School of Interactive Computing. MSUA is a one year program (three semesters), designed to give graduates a core of computing, urban planning, and data analysis and visualization skills to identify, analyze, and solve urban problems. Students will integrate those skills in an interdisciplinary way that other, single-discipline-oriented urban analytics degrees might not, and will experience a depth of urban problems addressable with data analytics. More information can be found at: https://planning.gatech.edu/master-science-urban-analytics.

LEHIGH UNIVERSITY

I-DISC
INSTITUTE FOR DATA, INTELLIGENT SYSTEMS, AND COMPUTATION

SUMMARY

I-DISC builds upon the foundation of Lehigh research expertise in areas such as machine learning, optimization, probabilistic modeling, data-driven decision making, high-performance & data-intensive computing, statistical signal and image processing, data representation & management, modeling & simulation, robotics & computer vision, business & management technology, and privacy & security.

DEGREES, PROGRAMS, SPECIALIZATIONS

I-DISC is a research institute and does not offer degrees. However, Lehigh offers the following related degrees:

- MS Data Science
- MS Business Analytics
- Undergraduate Minor in Data Science
- Certificates:
  - Business Analytics - Undergraduate
  - Data Science & Financial Analytics - Graduate
  - Data Analytics
  - Data Science (coming Fall 2022)

LOCATION

Building C, BC 229, Lehigh University Mountaintop Campus, 113 Research Drive Bethlehem, PA 18015

RESEARCH AREAS

The Institute for Data, Intelligent Systems, and Computation (I-DISC) brings together scholars from across Lehigh University with an interest in computationally focused research. I-DISC researchers build on expertise in fundamental computational approaches (such as machine learning, optimization, data science, and system architecture) and explore applied areas such as robotics, computer vision, finance, privacy and security, cyber-physical systems, energy, and materials science.
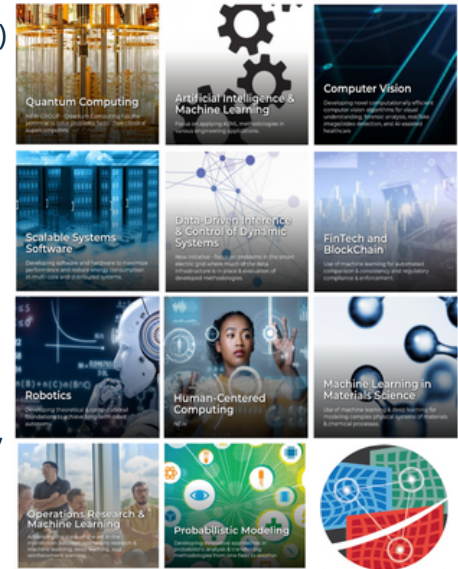
website: idisc.lehigh.edu/research

DATA AND COMPUTING EQUITY AND JUSTICE

I-DISC is committed to promoting equity and justice in data and computing through both:
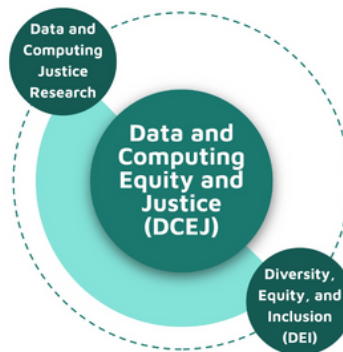
Data and Computing Justice Research: Scholarship related to identifying and mitigating the racist, sexist, ableist, homophobic, and other oppressive outcomes that are often produced by data and computation tools, as well as scholarship related to the positive outcomes that can emerge from careful design and implementation of such tools.

Diversity, Equity, and Inclusion (DEI): Policies, practices, and procedures for improving DEI, within I-DISC, within Lehigh as a whole, and within the broader data and computation community.

website: idisc.lehigh.edu/dcej

SUPPORT FOR FACULTY

I-DISC offers a number of opportunities and support to faculty members:

- I-DISC Fellow Services
- Faculty Support for Undergraduate Research
- Funding & Administrative Support for I-DISC Events at Lehigh
- Opportunities to Engage in High School Outreach Projects
- Grant Writing Services for Proposals

More information can be found on our website at idisc.lehigh.edu/opportunities-resources/faculty

EVENTS

I-DISC hosts and supports lectures, seminars, workshops and conferences throughout the year to further promote I-DISC's mission. These include faculty forums, coding retreats, grant-writing workshops, and networking events.

website: idisc.lehigh.edu/events

I-DISC FELLOWS

I-DISC Fellows are advanced Ph.D. students with expertise in machine learning, data science, optimization, social sciences, and other topics. Fellows provide short-term assistance to faculty members across the university to facilitate data- and computing-related aspects of their research.

Academic Data Science Alliance

College of Basic and Applied Sciences

MIDDLE TENNESSEE STATE UNIVERSITY

SUMMARY

Middle Tennessee State University offers Data Science degrees at all levels (BS, MS, and PhD) that emphasize the interdisciplinary nature of Data Science. MTSU also contains the Data Science Institute, which is an applied research focused entity that strives to help its partners solve complex problems.

DEGREES, PROGRAMS, SPECIALIZATIONS

- BS in Data Science
- MS in Data Science
- Ph.D. in Computational and Data Science
- Data Science Graduate Certificate

PROGRAM STATISTICS

- Number of Staff: 3
- Affiliated Faculty: 9
- Students enrolled in data science related programs (year 2022):
  - 52 BS, MS
  - 35 Ph.D.
  - 25 Grad. Certificate

SOCIAL

email: datascience@mtsu.edu
web: www.mtsu.edu/datascience
twitter: @mtsudatascience
linkedin:
www.linkedin.com/showcase/65268246/adminv
facebook:
www.facebook.com/mtsudatascience

LOCATION

1301 E Main St, Murfreesboro, TN 37132

PROGRAM OVERVIEW

The Data Science Bachelor's Degree at MTSU was announced in February of 2020. Even with the challenges of going remote, the program has exceeded enrollment expectations and has a vibrant community of students. This degree program combines classes from several different departments in two different colleges to give students the skills to use data to solve problems. Students gain a core set of knowledge and then choose one of three cognates to specialize in: Machine Learning; Business Intelligence and Analytics; and Inferential Thinking. There is also enough flexibility in the program that students can complete a minor in a field of interest to them. At the end of the program, students take a capstone course that combines all the knowledge they have gained into one project.

FOSTERING INCLUSIVE EXCELLENCE

MTSU started offering a graduate certificate in Data Science in 2020, and the Master's Degree in Data Science has been approved to start in Fall 2022.

The graduate certificate is an online program consisting of four courses with "Data Dive" events at the end of each course. During the Data Dives, students work in groups at an all-day event to use the course knowledge to analyze a data set.

The Master of Science in Data Science Program brings computer science, applied mathematics, and statistics together to prepare its graduates for seeking careers in data science or pursuing PhD in relevant areas. The curriculum contains a blend of theory and practice of computer science and applied mathematics in order to draw insights and to extract information from large data.

DATA AND OUR COMMUNITY

The existing Computational Science Ph.D. program at MTSU was recently renamed The Computational Science and Data Science (CDS) program. Computer models and computer simulations have become an essential part of the research repertoire. Going from application to computational results requires domain expertise, mathematical modeling, numerical analysis, algorithm development, software implementation, and the validation and visualization of results.

The increasing complexity of data sets has also become a significant challenge. Collecting and storing data has become easy, but the sheer volume of data produced across the disciplines has made traditional data analysis nearly impossible. The only way to survive this flood of data is to combine computer science, statistics, visualization, and numerical techniques into automated methods for discovering patterns within the numbers. Using domain expertise and cutting-edge techniques, we transform bits into understanding.

Because elements from several disciplines are utilized in the CDS program, integrating knowledge and methodologies differentiate this program from those with a singular focus. We use foundational areas such as computer science and numerical methods to enable cutting-edge research across disciplinary boundaries.

OUTREACH AND EDUCATION

The Data Science Institute at MTSU has the goal to take complex data and turn it into actionable information that adds value to an organization. The Institute is responsible outreach to industry leaders and to K-12 educators in the area. The Institute is also an entity for research in Data Science, bringing together research-focused faculty and staff with possibilities for collaborations across discipline boundaries.

NJIT | INSTITUTE FOR DATA SCIENCE

PROGRAM SUMMARY

In addition to the Institute for Data Science, NJIT established the Department of Data Science in 2021 as the newest addition to the Ying Wu College of Computing. The Department of Data Science offers undergraduate and graduate degrees, as well as graduate certificates, giving students the knowledge and skills needed to find success in the rewarding field of data science. To learn more about the programs offered please visit https://ds.njit.edu/programs.

PROGRAM STATISTICS

- Number of faculty: 6
- Number of technical staff: 2
- Affiliated Faculty: ~40

DEGREES, PROGRAMS, SPECIALIZATIONS

- B.S. in Data Science
- M.S. in Data Science
- Ph.D. in Data Science
- Certificate in Big Data Essentials
- Certificate in Data Mining
- Certificate in Data Visualization

SOCIAL

email: info.datascience@njit.edu
web: datascience.njit.edu

LOCATION

Newark, New Jersey

SUMMARY

Headed by Distinguished Professor David Bader, the Institute for Data Science focuses on cutting-edge interdisciplinary research and development in all areas pertinent to digital data. Beyond academic research, the institute interacts closely with the outside world to identify and solve important problems in the modern data-driven economy.

INSTITUTE OVERVIEW

The Institute houses four research centers: The Center for Big Data, The Cybersecurity Research Center, The Structural Analysis of Biomedical Ontologies Center (SABOC), and the newly established Center for AI Reseach.

Center for Big Data

Cybersecurity Research Center

Center of Artificial Intelligence

Structural Analysis of Biomedical Ontologies Center (SABOC)

NJIT
New Jersey Institute of Technology

Academic Data Science Alliance

NYU | Center for Data Science

INSTITUTE BIO

The Center for Data Science is the focal point for New York University's university-wide efforts in Data Science. CDS was established to advance NYU's goal of creating a world-leading Data Science training and research facility, and arming researchers and professionals with the tools to harness the power of Big Data.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Undergraduate Major in Data Science
- Undergraduate Joint Major in Computer Science and Data Science
- Undergraduate Joint Major in Data Science and Mathematics
- Undergraduate Minor in Data Science
- Master's in Data Science
- PhD in Data Science

CONTACT / SOCIAL

Program: datascience-group@nyu.edu
Twitter: @NYUDataScience
LinkedIn: www.linkedin.com/school/nyu-center-for-data-science
Facebook: @nyudatascience
Medium: nyudatascience.medium.com

LOCATION

60 5th Avenue
New York, NY 10011

ABOUT THE CENTER FOR DATA SCIENCE

The Center for Data Science (CDS) is the focal point for New York University's university-wide efforts in Data Science. The Center was established in 2013 to advance NYU's goal of creating a world-leading Data Science training and research facility, and arming researchers and professionals with the tools to harness the power of Big Data. Today, CDS counts 18 jointly appointed interdisciplinary faculty housed on three floors of our magnificent 60 5th Avenue building, one of New York City's historic properties. It is home to a top-ranked MS in Data Science program, one of the first PhD programs in Data Science, and a new undergraduate program in Data Science, as well as a lively Fellow and Postdoctoral program. It has over 60 associate and affiliate faculty from 25 departments in 9 schools and units. With cross-disciplinary research and innovative educational programs, CDS is shaping the new field of Data Science.
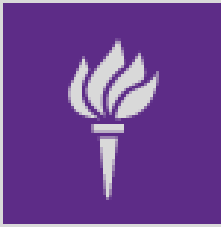
CDS RESEARCH

As CDS continues to establish itself as a leader in data science education and research, the center has focused on developing and expanding a number of academic research initiatives.

- The Center for Responsible AI (NYU R/AI) focuses on making "AI" synonymous with "responsible AI." The Center conducts research that engages in AI policy and regulation as well as teaching different constituents about AI and its social impact.
- Minds, Brains, and Machines is a joint effort by CDS, the Department of Psychology, the Center for Neural Science, and the Flatiron Institute. The research of MBM sits at the intersection of data science and cognitive science, leveraging progress in machine intelligence to advance our understanding of human intelligence.

  - The ML² Research Group consists of a team of researchers at NYU working on developing and studying state-of-the-art machine learning methods for natural language processing (NLP).
  - The Math and Data (MaD) group, a shared initiative, with CDS and the NYU Courant Institute, aims to advance the Mathematical and Statistical foundations of Data Sciences, specializing in signal processing and inverse problems, machine learning and deep learning, and high-dimensional statistics and probability.
  - The STAT group (Statistics: Tools, Algorithms, and Theory), a joint effort with CDS and the NYU Courant Institute, seeks to advance the state-of-the-art in statistics, by developing new methodological, computational, and mathematical approaches to statistical problems and to their applications in data science and machine learning.

NORTHWESTERN INSTITUTE ON COMPLEX SYSTEMS

SUMMARY

NICO's mission is to incubate innovative collaborations that leverage complexity, networks, and data science to address societal challenges. Faculty initiatives nurtured by NICO are transforming areas as disparate as synthetic and quantitative biology, sustainability and resilience engineering, computational social sciences, business, and the law.

PROGRAM STATISTICS

- Number of professional staff: 7
- Affiliated faculty: 83
- Number of students in undergraduate programs:
  - 375 students enrolled in minor
  - 138 enrolled in major

THE DATA SCIENCE REVOLUTION

We live in a world where not just the amount, but also the diversity and complexity of digital information continues to grow exponentially. Materials simulations and astronomy images are pushing the boundaries of exploration. Social networks enable the exchange of information between billions of individuals; medical devices, GPS devices and bar code scanners allow the exchange of information between machines.

Importantly, "bigger is not just better, it is different:" This is what researchers and entrepreneurs across the academic, government, and private sectors realize as the Data Science revolution continues to unfold and affect our scientific inquiry process. Figuring out how to efficiently, effectively, and reliably extract new knowledge from big and/or complex datasets presents us with new algorithmic and practical challenges, but also enables us to ask new questions.

NORTHWESTERN'S DATA SCIENCE INITIATIVE

At Northwestern, we address Data Science from a different perspective. We contend that Data Science is not about the absolute size of the data but about a change in scale. Indeed, we believe that all knowledge creation fields are being transformed by the relative increase in scale of the data available to scholars.

Data science is affecting every aspect of Northwestern's learning and research enterprises. The Data Science Initiative was established in 2015 and is managed by the Northwestern Institute on Complex Systems (NICO).

LOCATION

600 Foster Street
Evanston, IL 60208

SOCIAL
email: nico@northwestern.edu
web: datascience.northwestern.edu
twitter: @NICOatNU
linkedin: www.linkedin.com/groups/6615433
facebook: www.facebook.com/NICONorthwestern
youtube: bit.ly/NICOatNUonYouTube

DEGREES, PROGRAMS, AND SPECIALIZATIONS

Undergraduate degrees:

- Major in Data Science (offered by Department of Statistics): statistics.northwestern.edu/undergraduate/data_science_major/index.html
- Minor in Data Science (offered by Department of Statistics): statistics.northwestern.edu/undergraduate/data_science_minor/index.html
- Data Science and Engineering Minor (offered by School of Engineering): www.mccormick.northwestern.edu/data-science-engineering

Master degrees:

- *Master in Data Science (offered by School of Professional Studies): sps.northwestern.edu/masters/data-science*
- *Master of Science in Analytics (offered by School of Engineering): www.mccormick.northwestern.edu/analytics*
- *Master of Science in Artificial Intelligence (offered by School of Engineering): www.mccormick.northwestern.edu/artificial-intelligence*

Data Science Certificates:

- *Integrated Data Science Certificate (offered by CIERA): ciera.northwestern.edu/2016/01/01/integrated-data-science-certificate-in-development-at-ciera*
- *Health Sciences Integrated Program (offered by School of Medicine): www.feinberg.northwestern.edu/sites/hsip*
- *Analytics and Modeling (offered by School of Professional Studies): catalogs.northwestern.edu/sps/certificates/graduate/data-science/analytics-modeling-graduate-certificate*
- *Analytics Management (offered by School of Professional Studies): catalogs.northwestern.edu/sps/certificates/graduate/data-science/analytics-management-graduate-certificate*
- *Data Engineering (offered by School of Professional Studies): catalogs.northwestern.edu/sps/certificates/graduate/data-science/data-engineering-graduate-certificate*
- *Artificial Intelligence (offered by School of Professional Studies): catalogs.northwestern.edu/sps/certificates/graduate/data-science/technology-entrepreneurship-graduate-certificate*

FACULTY RESEARCH NETWORKING LUNCHEONS

Interdisciplinary groups of faculty meet monthly during the academic year for discussion and exploration of significant and upcoming research areas with a strong data science and analysis focus. Overall goals of this "interdisciplinary research networking" model include new faculty connections and collaborations, innovative research agendas, collaborative publications, research grant proposals, potential new and joint hires, and identification of funding opportunities and new projects.

Data Science Academy

NC STATE UNIVERSITY

SUMMARY

NC State launched the Data Science Academy (DSA) in July 2021 to build data science efforts across all university colleges and to address all pillars of its mission: research, teaching and outreach to the state of N.C. **Rachel Levy,** professor of mathematics, is the DSA's Executive Director.

DEGREES, PROGRAMS, SPECIALIZATIONS

300 students enrolled in 14 1-credit, project-based course sections in AY2021-2022.

The DSA is developing data science certificates and an undergraduate minor.

Graduate and undergraduate units have integrated DSA courses into their programs of study.

SOCIAL

email: datascienceacademy@ncsu.edu
web: datascienceacademy.ncsu.edu
twitter: @NCStateDSA

LOCATION

NC State Data Science Academy
2 Broughton Drive
Campus Box 7111
Raleigh, NC 27695-0001

DATA SCIENCE AT NC STATE

When the DSA launched, Chancellor Randy Woodson noted "Data science is integrated in subject matter and research programs across all 10 NC State colleges and is part of our 'think and do' mindset. At NC State, we believe data science is for everyone."

The DSA is part of strengthening interdisciplinary connections between data science initiatives at NC State, as illustrated below.
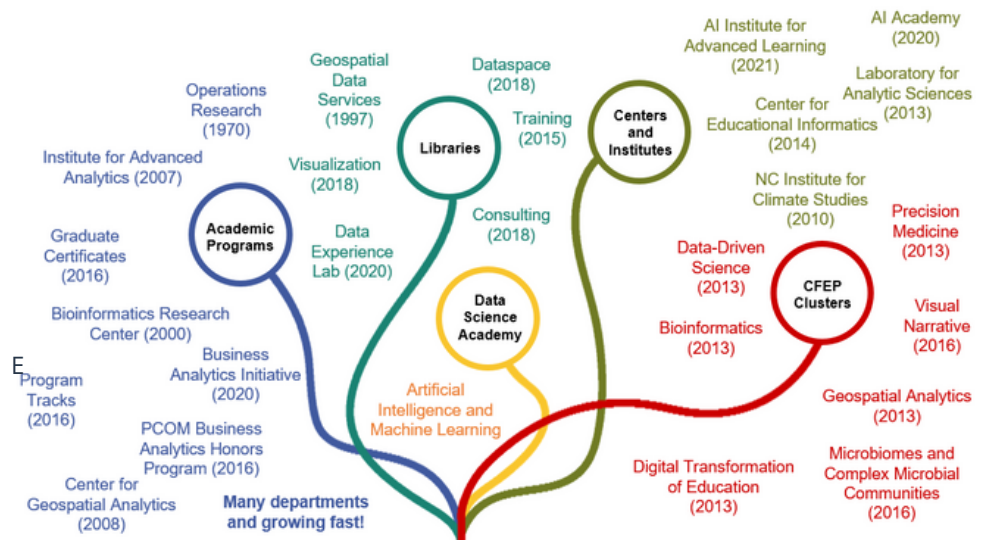
RESEARCH ENABLEMENT

The DSA collaborates with the NC State Office of Research Innovation to develop interdisciplinary collaborations across campus. Seed grants and virtual think tanks spark new conversations and collaborations.

INDUSTRY PARTNERSHIPS

The DSA works with the NC State Office of Partnerships to build connections with local, national and global industrial partners in research collaborations and workforce development.

AGRICULTURAL EXTENSION

Farmers and extension agents across North Carolina are using data every day to make critical decisions about all aspects of agriculture. The DSA supports development of their data science expertise.

PK-12 OUTREACH

In collaboration with the NC State Science House and Friday Institute, the DSA is building data science opportunities for students and teachers across North Carolina and beyond.

Data Science Academy

NC STATE UNIVERSITY

FALL 2021 EVENTS

Data science career panel

As part of the DSA virtual career expo, the DSA organized an evening panel with co-sponsorship from the National Consortium for Data Science. Diverse panelists shared their data science career paths and experiences.

Tackling Climate Change with Machine Learning

Priya L. Donti, co-founder and chair of Climate Change AI, presented a public lecture on what machine learning can, and can't, do to tackle climate change.

TEACHING SPOTLIGHT

*Postdoctoral fellow **James Harr**'s Data Visualization class presents final projects in the 360-degree Visualization Studio in the D.H. Hill Jr. Library*

DATA ANALYTICS RESEARCH CONSULTING

The DSA collaborates with the NC State University Libraries to provide analytics consulting.

Emily Griffith, DSA Director of Analytics Consulting and associate research professor of statistics, leads an interdisciplinary corps of graduate research assistants to provide data science expertise across campus.
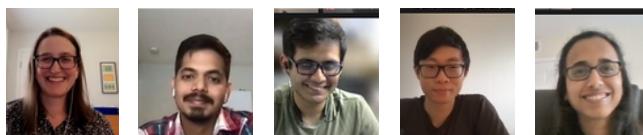
1-CREDIT PROJECT-BASED COURSES

Fall 2021
- R/Python for Data Science
- Exploratory Data Analysis for Big Data
- Data Wrangling and Web Scraping
- Data Science for Policy
- Data Visualization

Spring 2022
- R/Python for Data Science (*3 sections*)
- Exploratory Data Analysis for Big Data
- Natural Language Processing
- Data Wrangling and Web Scraping
- Data and Ethics
- Data Physicalization
- Machine Learning for Practitioners

At NC STATE, data science is for everyone!

Academic Data Science Alliance

THE OHIO STATE UNIVERSITY
TRANSLATIONAL DATA ANALYTICS INSTITUTE

SUMMARY

The Translational Data Analytics Institute is a community of researchers at the forefront of interdisciplinary, data-enabled science, scholarship and creative expression with an emphasis on significant societal impact. Its 200+ core faculty and affiliates collaborate across 50+ disciplines to innovate data science and analytics solution for the greater good.

PROGRAM STATISTICS

- Number of faculty: 40
- 7 current + 7 actively/near-term recruiting…and growing. (+ 4 for newly funded NSF Harnessing the Data Revolution: Imageomics Institute)
- Affiliated researchers: >250
- Students: 43 in the MTDAI program

DEGREES, PROGRAMS, SPECIALIZATIONS

- Master in Translational Data Analytics (MTDA)
- (Note: Other data science-related academic programs exist at Ohio State, but the MTDA was designed by and is administered by TDAI.)

LOCATION

1760 Neil Ave
Columbus, OH 43210

OUR HISTORY

The Translational Data Analytics Institute (TDAI) was established at Ohio State in 2015 as a foundational component of the university's $500 million Discovery Themes initiative to address grand challenges—a recognition of big data's significance in answering the most pressing questions in science and solving the most vexing problems of the human condition. The institute was designed to foster the interdisciplinary study, development and application of data science and analytics solutions among its 40 core faculty, 250+ affiliates and students across Ohio State's 15 colleges. In 2019, TDAI opened a 21,000-square-foot home in Pomerene Hall, the result of a $125 million joint investment by the university and state of Ohio to create a regional hub of data science and analytics collaboration and excellence.

OUR RESEARCH AND SCHOLARSHIP

Although collective research and scholarship interests of our faculty affiliates are quite broad and diverse, the Institute has focused on 5 strategic directions: Foundations of AI; Smart Mobility; Responsible Data Science; AI & Health; Environment & Sustainability. The value that the Institute brings to our faculty-led research efforts in these strategic areas is the ability to facilitate internal and external partnership-building across numerous disciplines. We do this by: awards of funding and in-kind resources for research pilots, extramural proposal development, and early-stage idea exploration; data science consulting services; postdoctoral fellowships; seminar series; research infrastructure; facilitated ideation sessions; and data science and analytics workshops to support training of students and postdocs on faculty-led research teams. Our approach has proven to be successful, as evidenced in part by: promotion and tenure of

many of our core faculty; increased submissions of large, interdisciplinary research proposals to extramural funding agencies; and the recent award of 3 NSF-funded institutes: AI Institutes in cyberinfrastructure and edge computing, and an HDR Institute in a new field of imageomics.
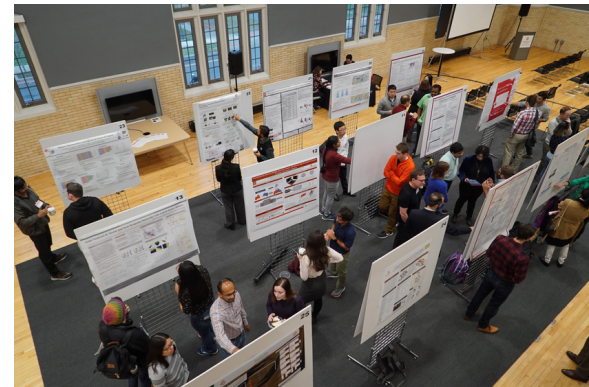
OUR TRAINING AND EDUCATION PROGRAMS

In 2020 TDAI launched a master's degree program in translational data analytics to upskill and reskill working professionals. This program is a partnership between TDAI, the Graduate School, Advanced Computing Center for Art & Design, and the departments of statistics, design, and computer science and engineering. The MTDA integrates design thinking into foundations of data science, programming, machine learning, and data visualization, uniquely positioning students to be master data storytellers.

TDAI also provides non-traditional training opportunities by (1) our Data Literacy workshops in which participants learn the basics of data wrangling and coding, and (2) employing students in our Data Science Consulting Services. TDAI has partnered with OSU Libraries to lead the delivery of the Carpentries at Ohio State. The Carpentries are an international organization that delivers a collection of lessons aimed at introducing data and coding skills to spread data and programming literacy. As part of the membership, several graduate students from TDAI faculty research teams have become certified Carpentries instructors and are now teaching TDAI's Data Literacy workshops.

OUR TRAINING AND EDUCATION PROGRAMS

TDAI's Data Science Summer Camp serves as an exemplar of the Institute's intentional efforts to broaden participation in STEM fields. The camp is a week-long program designed for middle and high school students to increase exposure to data and analytics earlier in their education using a fun, hands-on approach to learning. Collectively, the camp serves students from Columbus City and Metro school districts in both Franklin and

Delaware counties. Now in its fourth year, the camp has expanded to 83 campers representing 42 unique middle and high schools across the state of Ohio. With an emphasis on recruiting underrepresented minorities (URMs), more than 80% of the campers in 2021 self-identified as URM – the fourth straight year that we have achieved a URM enrollment of >68%.

SOCIAL
email: tdai@osu.edu
web: tdai.osu.edu
twitter: @OSUbigdata
linkedin:
www.linkedin.com/company/tdai
nstitute

Penn State Data Science

PENNSYLVANIA STATE UNIVERSITY

SUMMARY

Penn State's diverse Data Science portfolio includes research and educational programs at many of the University's 24 campuses and 16 academic colleges. Academic units offer degree programs aligned with their disciplinary focus, and the interdisciplinary Institute for Computational and Data Sciences provides University-wide data sciences research support.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Bachelor of Science in Data Sciences, with options in Statistical Modeling, Applied Data Sciences, and Computational Data Sciences
- Bachelor of Science in Social Data Analytics
- Master of Informatics, Data Sciences Concentration
- Master of Science in Data Analytics
- Master of Science in Spatial Data Science
- Master of Professional Studies in Data Analytics
- Master of Professional Studies in Artificial Intelligence
- Dual-Title Doctor of Philosophy in Social Data Analytics
- Graduate Certificate in Data Analytics
- Undergraduate Certificate in Geospatial Big Data Analytics

LOCATION

201 Old Main, University Park, Pennsylvania 16802

STUDENT ENGAGEMENT

Home to nearly 100,000 students, Penn State offers numerous opportunities for students to hone their data science skills and build their resumes. Students can network with peers by joining Nittany Data Labs, make professional connections by joining a handful of professional organization chapters, including the Association of Women in Computing and the Institute of Electrical and Electronics Engineers, build their coding and programming skills through the 24-hour HackPSU marathon, or become entrepreneurs and change-makers in the industry-funded Nittany AI Alliance competition

(photo credit: Penn State)

INTERDISCIPLINARY RESEARCH

An R1 university, Penn State has nearly $1 billion in research expenditures annually, and many of these projects rely on data science. Penn State's Institute for Computational and Data Sciences (ICDS), part of the Office of the Senior Vice President for Research, supports data science broadly across the University. ICDS offers

funding; organizes faculty from different disciplines around major funding opportunities; operates a supercomputer, known as Roar, for the Penn State community; staffs a team of data science support consultants who collaborate with faculty in dozens of disciplines; and hosts university-wide events around data science topics such as reproducibility, bias in algorithms, and more. In addition to ICDS, several federally or university-funded research centers gather researchers around foci such as artificial intelligence for engineered systems, or trustworthy machine learning.

(photo credit: Geoffrey King / Penn State)

CONTACT INFORMATION FOR DATA SCIENCE PROGRAMS

Eberly College of Science / Department of Statistics
• Contact: 814-865-1348 or stat-advising@psu.edu

College of Engineering / Department of Computer Science and Engineering
• Contact for B.S. in Data Sciences, Computational Data Sciences Option: 814-865-9505 or arc88@psu.edu

College of Information Sciences and Technology
• Contact for B.S. in Data Sciences, Applied Data Sciences Option: 888-225-8707 or futurestudents@ist.psu.edu
• Contact for Master of Informatics Program: 814-863-0591 or istgradprograms@psu.edu

College of the Liberal Arts / Department of Political Science
• Contact for B.S. in Social Data Analytics: 814-863-8346
• Contact for dual-title Ph.D. in Social Data Analytics: bdesmarais@psu.edu

College of Earth and Mineral Sciences / Department of Geography
• Contact for Certificate in Geospatial Big Data Analytics: 814-865-3433 or geography@psu.edu

Penn State Great Valley / School of Graduate Professional Studies
• Contact for Artificial Intelligence Program (offered online through Penn State World Campus): ArtificialIntelligence@psu.edu
• Contact for M.S. and M.P.S. in Data Analytics: gvengin@psu.edu or 610-648-3243

Texas A&M Institute of Data Science

TEXAS A&M UNIVERSITY

SUMMARY

The Texas A&M Institute of Data Science (TAMIDS) pursues new approaches to Data Science research, education, operations and partnership. These approaches cross college boundaries to connect elements of Data Science from engineering, technology, science and the humanities, and inform wider social challenges.

PROGRAM STATISTICS

- Number of staff: 19
- Affiliated Faculty: 216
- Students: 22

DEGREES, PROGRAMS, SPECIALIZATIONS

- MS in Data Science
- Certificate in Data Analytics for the Petroleum Industry
- Thematic Data Science Labs
- Professional Education
- Capstone Industry Mentorship
- Hackathon Sponsorship
- Visiting Researcher Program

SOCIAL

email: duffieldng@tamu.edu
web: tamids.tamu.edu
linkedin:
www.linkedin.com/company/tamudatascience
facebook: www.facebook.com/TamuDataScience

LOCATION

Texas A&M Institute of Data Science
John R. Blocker Building, Suite 227
155 Ireland Street, TAMU 3156
College Station, TX 77843-3156

EDUCATION

The Master of Science in Data Science degree is an on-campus interdisciplinary program with tracks offered by the Texas A&M Departments of Computer Science and Engineering, Electrical and Computer Engineering, Mathematics, and Statistics, with coordination across departments by TAMIDS. The Undergraduate Certificate in Data Analytics for the Petroleum Industry draws students from the Texas A&M Colleges of Engineering, Science, Geosciences, and the Mays Business School, and is managed by TAMIDS and the Department of Petroleum Engineering. Our programs follow the Curricular Paths framework that helps students from diverse disciplines prepare to work in interdisciplinary teams, through Capstone and Practicum courses.

TRAINING AND WORKFORCE DEVELOPMENT

TAMIDS Training offers bootcamps, webinars, and summer courses to help students, researchers, and faculty develop skills through hands-on experience in Data Science. The Professional Education Workshops in Data Science Foundations and Computational Practice help technical specialists extend their expertise to encompass the systems, methods, and tools of Data Science.

THEMATIC DATA SCIENCE LABS

TAMIDS Thematic Labs develop knowledge, resources, and community around emerging areas of Data Science, encompassing research, education, and outreach. Each lab pilots new interdisciplinary research, develops for-credit courses, training, and case studies, and hosts a program of seminars and workshops. The Scientific Machine Leaning Lab combines Scientific Computation and ML to incorporate scientific model constraints in ML algorithms and enable prediction of performance of complex multiscale, multiphysics systems using sparse, low-fidelity, and heterogeneous data. The Data Justice Lab brings together computational data scientists and social justice-oriented social scientists to use and study the impact of Data Science in social domains such as health, education, the built environment, crisis management, and economics. The VIVID Lab for Visceral Intersensory Visualiztion & Information Design takes visualization beyond the traditional notion of a visual display to encompass interaction with information through sensors, AR and VR, craft, and fabrication. The Operational Data Science Lab works with Texas A&M infrastructure and administration to capitalize on institutional data investments, improve campus operations, and engage Texas A&M researchers in working on operational problems.

PARTNERSHIP AND ENGAGEMENT OPPORTUNITIES

TAMIDS partners with industry for student engagements including sponsorship of capstone projects, hackathons, and competitions, and develops certificate programs preparatory for emerging industry needs. TAMIDS matches our faculty affiliates to industrial opportunities. The TAMIDS visitor program supports researchers from universities, research labs and industry to collaborate with TAMIDS and the broader TAMU Data Science community.

Academic Data Science Alliance

Data Intensive Studies Center

SUMMARY

The Data Intensive Studies Center at Tufts is a cross-university center focused on enabling data-intensive research, scholarship and education.

DEGREES, PROGRAMS, SPECIALIZATIONS AT TUFTS

- Data Science M.S.
- Data Analytics M.S.
- Data Science Certificate
- Data Analytics Certificate
- Specializations in: Biomedical statistics, bioinformatics, Bayesian statistics, algorithmic fairness, machine learning, agent based modeling, uncertainty quantification, scientific computing, computational mathematics.

PROGRAM STATISTICS

- Number of faculty: 3
- Number of staff: 5
- Affiliated Faculty: 6
- Students: 100

SOCIAL

email: disc@tufts.edu
web: disc.tufts.edu

LOCATION

Data Intensive Studies Center
Tufts University
Joyce Cummings Center
177 College Ave.
3rd Floor Suite 336
Medford, MA 02155

PROGRAM OVERVIEW

The Data Intensive Studies Center serves idea as a vital hub that cultivates collaborations and fosters a culture where computational, data and statistical sciences integrate with domain sciences. The center provides a clear framework that brings together an array of disciplines, working across existing departments, schools, and campuses to facilitate cutting-edge research and to prepare Tufts

graduates for a future in which data sciences are increasingly important. As a nexus, the Center also hosts various activities, including forums, mini-symposia, and seminars.

TRAINING AND WORKSHOPS

The Data Intensive Studies Center offers training workshops in data and computing sciences with contributors from several academic units including Tufts Technology Services and Tufts Clinical and Translational Science Institute and the Tufts Data Lab. Short 3-hour workshops as well as multi day 15-hour workshops have been offered on topics such as Bayesian Methods, Bioinformatics, Data Visualization, and Introductory Python Programming. Workshops are open to all students, faculty and staff.

DISCIPLINARY PARTNERSHIPS

Data sciences can enable transformative research in many disciplines. However, initiating such work can be difficult and needs partners and expertise that may be difficult to find. The Data Intensive Studies Center helps Tufts researchers overcome these barriers. Through this effort we seek to accelerate well defined X+Data Sciences efforts on campus that can target high impact research and in appropriate areas large sponsored research efforts e.g. NSF TRIPODS Phase II, NIH BIG Data to Knowledge, AI Institutes. DISC will supports such efforts with a seed grant of up to $20,000 and access to data scientists, data sets and other expertise at DISC. Through these projects DISC engages faculty, staff, and students in trans-disciplinary research.

PIONEERING RESEARCH
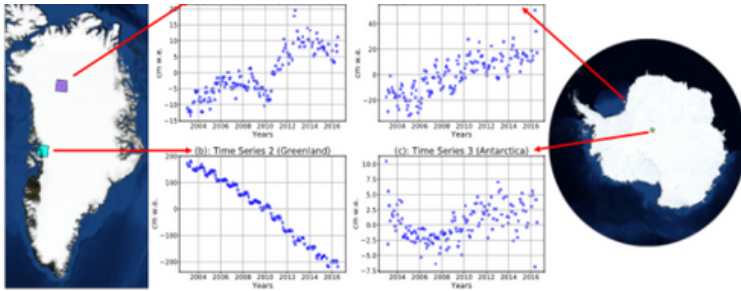
Hierarchical Learning for Noisy datasets in ice sheet remote sensing, (*Patra and Shekhar*)
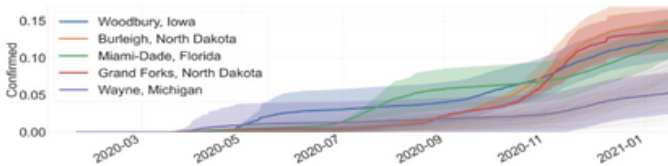
New large deviation principle based anomaly detection in COVID data (*Patra and Guggilam*)
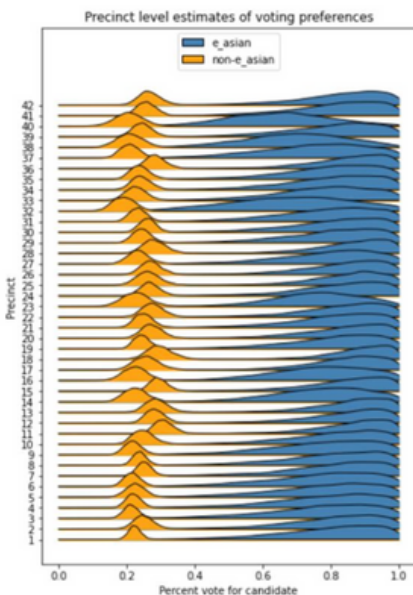
Racially polarized voting using ecological inference (*Knudson with Duchin group*)
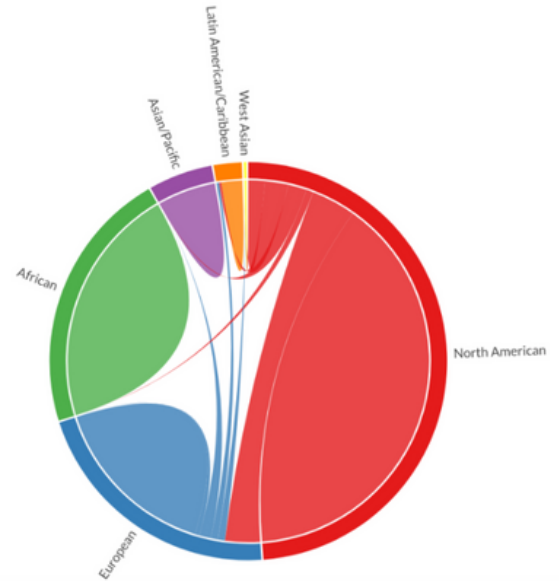
Tufts art collection - diversity measures - Haensch and SMFA - art hackathon planned in the Fall (*Haensc and Deitsch*)
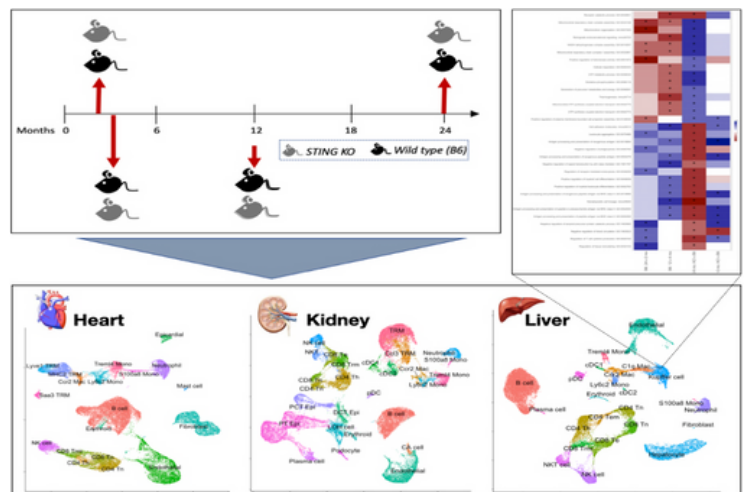
Changes to the immunologic activity landscape are implicated in a variety of age-associated deleterious health outcomes - effect of inflammatory cytokine activating protein, Stimulator of interferon genes (STING). Gene set-projection methods to characterize biological pathways with age-associated changes. (*E. Reed, A. Tai and Shruti Sharma*)

Academic Data Science Alliance

UNIVERSITY OF AMSTERDAM
Data Science Centre

SUMMARY

The University of Amsterdam Data Science Centre mission is to enhance the university's research by developing, sharing and applying data science methods and technologies. As a coordinating hub within the UvA Library, the centre is uniquely positioned to facilitate knowledge exchange as well as training in data-driven research.

DEGREES, PROGRAMS, & SPECIALIZATIONS

The Data Science Center offers 10 Bachelors and 14 Masters degrees related to Data Science. These include:

- Masters in Informational Studies: Data Science
- Masters in Behavioral Data Science
- MBA in Big Data & Business Analytics

PROGRAM STATISTICS

- Affiliated Faculty & Staff: 18
- Students enrolled in data science related programs (year 2020): 2743 students
  - 1450 bachelor's students
  - 1293 master's students

LOCATION

UvA University Library
Singel 425, 1012 WP
Amsterdam
The Netherlands

PROGRAM OVERVIEW

The UvA's Data Science Centre is focused on accelerating data driven research across the university's entire research landscape.

To do so, the centre is embedding data scientists and engineers throughout the university directly in research groups. These affiliated data scientists and engineers share knowledge and experience through weekly sessions at the library - the intellectual crossroads of the university.

Our current Affiliated Faculty will be growing from 18 to 35 members by 2025.

DATA SCIENCE ACADEMY

Continuing to develop at an academic level and boost your career at the same time? UvA Academy makes the most recent insights from science and practice accessible to professionals. Our short-term courses, masterclasses and webinars help you deal with issues of today and tomorrow.

Classes and webinar topics include:
- Analytics Translator
- Artificial Intelligence for Managers
- Applied blockchain
- Data Analytics & Quantitative Trading
- Data Science Essentials
- Data Science for Auditors
- Effectief Toezicht op Artificial Intelligence
- Ethics in Information Design
- Fintech: Blockchain, Cryptocurrencies & Smart Contracts
- Machine Learning & AI in Finance

Find the full list of topics here

UNIVERSITY OF AMSTERDAM
Data Science Centre

GRADUATE RESEARCH OPPORTUNITIES

The University of Amsterdam has a strong foundation of cutting edge research in artificial intelligence and data science. Building on this foundation, the DSC supports 7 PhD positions that are performing research into new data science methods that help to tackle hard challenging problems in a given domain.

The 7 positions include all seven faculties. Examples of the projects include:

- The GPU Dentist: Instilling Domain Knowledge in Deep Networks through Hyperbolic Geometry
- Building better vision models using pre-cortical inductive biases
- Natural Language Processing and Responsible Data Management for Mental Health Research

https://dsc.uva.nl/content/news/2021/10/seven-new-interdisciplinary-data-science-phd-positions.html

AMSTERDAM DATA SCIENCE

The University of Amsterdam Data Science Centre is proud to be partner of the Amsterdam Data Science (https://amsterdamdatascience.nl) a network organization that brings together industry and academia to help facilitate the city's ecosystem. These includes the four large knowledge institutions in Amsterdam and 20 other partners. Its meetup network has over 8000 members. Amsterdam Data Science also organizes thesis fairs introducing students to institutions and companies where they can do their master's thesis.

SOCIAL
email: dsc@uva.nl
web: dsc.uva.nl
twitter: @uva_dsc

Academic
Data Science
Alliance

Data Science at the University of Arizona

UNIVERSITY OF ARIZONA

SUMMARY

The University of Arizona is a globally recognized research leader. The Data Science Institute enables research collaborations between faculty, specialists, and students and expands the base for experiential learning and research innovations. Data Sciences Academy provides academic programming and K-12 Teacher Development for early exposure to data science concepts and skills. The Institute for the Future of Data and Computing compliments these activities through coordinated approaches and focusing on workforce development.

PROGRAM STATISTICS

- Number of technical staff: 14
- Affiliated Faculty: 86

DEGREES, PROGRAMS, SPECIALIZATIONS

The University of Arizona offers multiple undergraduate and graduate degree programs including:

- Statistics & Data Science Undergraduate Degree
- Computer Science Undergraduate Degree
- Information Science Undergraduate Degree
- Management Information Systems Undergraduate Degree
- Systems and Industrial Engineering Undergraduate Degree
- Population Health Data Science Minor
- And more

SOCIAL

email: rii-datascienceinstitute@arizona.edu
web: datascience.arizona.edu

LOCATION

1230 N. Cherry Avenue
Tucson, AZ 85721

DATA SCIENCE INSTITUTE

The Data Science Institute facilitates collaboration across an increasingly diverse and active Data Science community by providing workforce development, essential technological assistance, and training to University partners. Formerly Data7, the Data Science Institute (DSI) aims to foster the next generation of data-driven research by encouraging university-wide interdisciplinary collaboration, gaining external visibility, developing industry alliances, and increasing funding for UA research. By connecting UA researchers and aligning institutional expertise, computational resources, and infrastructure, DSI enables investigators to ask more complex questions and achieve outcomes not easily attainable as individual investigators or within purely disciplinary teams. DSI provides initial support to University of Arizona research projects by funding part-time graduate students, postdoctoral fellows, technical staff, and computational infrastructure – all working to start new collaborations or broaden existing collaborations.

DATA SCIENCES ACADEMY

Undergraduate students can find a spectrum of opportunities for the Data Sciences at the University of Arizona where they can choose a Data Science as both a major or a minor. The Data Sciences Academy initiative funds personnel and operations to support a university-wide endeavor to coordinate and lead in all areas of the data sciences and all aspects of university research, education and outreach.

INSTITUTE FOR THE FUTURE OF DATA AND COMPUTING (IFDC)

The IFDC goal is to ensure that Arizona is American's leader in computer and data innovation. IFDC is a new transdisciplinary institute being built to usher in the next generation of data, networking, and computing platforms. It provides seed funding to early-state faculty with interests in research with industry links and in mentoring students in company internships. IFDC creates and implements trainings for the University community and community and regional partners to engage with cutting-edge applications of data and computing. It creates a corporate engagement landscape of internships and other experiential learning opportunities that support students in those experiences.

CYVERSE - THE OPEN SCIENCE WORKSPACE FOR COLLABORATIVE DATA-DRIVEN DISCOVERY

CyVerse strives to create an innovative, comprehensive, generic, and foundational cyberinfrastructure in support of life science research. CyVerse develops cyberinfrastructure that uniquely enables scientists across the diverse fields that comprise life sciences to address Grand Challenge questions in new ways, to stimulate and facilitate cross-disciplinary research, to promote biology and computer science research interactions, and to train the next generation of scientists on the use of cyberinfrastructure in research and education

Academic Data Science Alliance

BIDS.
BERKELEY INSTITUTE FOR DATA SCIENCE

BERKELEY INSTITUTE FOR DATA SCIENCE

UNIVERSITY OF CALIFORNIA, BERKELEY

SUMMARY

The Berkeley Institute for Data Science (BIDS) facilitates innovative data-intensive research and open source software development across an increasingly diverse data science community of domain and methodological experts from the University of California, Berkeley, and beyond.

DEGREES, PROGRAMS, SPECIALIZATIONS

Interdisciplinary Research; Open Source Research Software; and Data Science Training Programs, including the BIDS Cross Domain Initiatives (XDs), research fellowships and internships, cross-disciplinary discussion groups, public lectures, and training seminars/webinars.

PROGAM STATISTICS

Number of staff: 10
Number of affiliates: 70

LOCATION

University of California
Berkeley, California

BIDS.BERKELEY.EDU

BIDS IS A CENTRAL HUB OF DATA-INTENSIVE RESEARCH, OPEN SOURCE RESEARCH SOFTWARE, AND DATA SCIENCE TRAINING AT UC BERKELEY.

BIDS programs and initiatives are designed to facilitate collaboration across an increasingly diverse and active data science community of domain experts – from the life, social, and physical sciences, and the humanities – as well as methodological experts from computer science, statistics, and applied mathematics. Since its launch in 2013, BIDS has cultivated an environment of open inquiry and discovery for data-intensive research, and we continue to seek new and creative ways to cross traditional academic boundaries and engage a diverse community of researchers representing a wide array of disciplines.

INTERDISCIPLINARY RESEARCH

BIDS supports innovative cross-disciplinary research for faculty, staff, postdoctoral fellows, and graduate and undergraduate students in a variety of fields across the life, health, social, and physical sciences, the humanities, and in computer science, statistics, and applied mathematics.

OPEN SOURCE RESEARCH SOFTWARE

BIDS contributes to open source, open science, and open data resource communities in support of academic research.

DATA SCIENCE TRAINING PROGRAMS

BIDS' data science training programs facilitate innovative collaboration across all research disciplines, providing a variety of opportunities for researchers at all levels, including the following:

BIDS Cross Domain Initiatives (XDs)
are cross-disciplinary research communities working together to identify common principles, algorithms and tools to advance research; to break down boundaries between domains; and to facilitate new collaborations among like-minded researchers.

BIDS Fellowship and Internship Programs
provide opportunities to cross interdisciplinary boundaries in a range of research fields across all domain areas. The **BIDS Data Science Fellowship** is a 2-year research training program open to postdocs and graduate students who are dedicated to advancing data science and undertaking innovative, cross-disciplinary data-intensive research. The **Berkeley Computational Social Science Training Program** recruits predoctoral students representing a variety of degree programs and expertise areas in the social sciences, including demography, public health, public policy, epidemiology, social welfare, and sociology. Participating fellows develop advanced computational and data science analytics skills to address urgent needs in biomedical, behavioral, social and clinical research. The *Innovate For Health* initiative is a collaboration between BIDS at UC Berkeley, the Bakar Computational Health Sciences Institute at UCSF, and Janssen Research & Development, who together launched the **Data Science Health Innovation Fellowships** to recruit data scientists for high-impact, data-driven healthcare research. This program provides 2 years of financial support to develop and execute health-related translational research that optimizes data-driven discovery and addresses currently unmet patient needs.

EDUCATION AND OUTREACH

BIDS' education and outreach programming endeavors to engage a wide variety of audiences with events ranging from discussion groups and workshops to training seminars, webinars, and public lecture series. BIDS has created and actively maintains a reputation of open engagement, which has been an integral factor in fostering collaboration and data science engagement with research.

BIDS.BERKELEY.EDU

International Computer Science Institute

SUMMARY

The International Computer Science Institute (ICSI), a non-profit institute in Berkeley, CA, conducts pioneering research in computer science and data science, focusing on areas such as big data, ML, NLP, AI, networking, security, and cloud computing. ICSI values and encourages scientific excellence, trans-disciplinarity, and social impact, from open-source initiatives to successful startups. The Institute welcomes researchers from around the world and across sectors who strengthen our research with diverse expertise, perspectives, and backgrounds.

PROGRAM STATISTICS

- Number of faculty: 9
- Number of research staff: 26
- Affiliated researchers: 32
- Students: 22

DEGREES, PROGRAMS, SPECIALIZATIONS

- Research in the following areas: Data Mining and Statistical Analysis, Big Data, ML, NLP, AI, Computer Vision, Audio & Multimedia, Privacy, Security, Cloud Computing, HCI, CSCW, Data Science Studies

LOCATION

2150 Shattuck Ave, Suite 250, Berkeley, California 94704

RESEARCH WITHOUT BORDERS

As one of the few independent, nonprofit research institutes in the United States, ICSI is a place where academia, government, industry, and social impact organizations come together to inspire world-class research and ground-breaking developments. ICSI welcomes researchers in any computationally-related field, from any country, to work with established staff researchers, UC Berkeley professors, and their networks of academic, government, and industrial partners.

Our research covers many topics, but here are three areas where ICSI has had recent impact:

- **Big Data and ML:** ICSI covers a range of topics in both the foundations and applications of large-scale data analysis, including stochastic convex and non-convex optimization, randomized linear algebra, numerically intensive machine learning, parallel and distributed computing, as well as a range of industrial and scientific applications. In recent years, we have focused on practical theory for neural network generalization, dynamical systems aspects of scientific machine learning, and developing principled methods to couple data-driven machine learning models and domain-driven scientific models.

- **Networked Systems:** In recent years we have helped develop Software-Defined Networking and Lean NFV, and our most recent efforts involve proposals for reshaping the cellular infrastructure, evolving the Internet architecture, and radically reimagining the cloud computing ecosystem. Overall, our research in networked systems covers a wide range of topics, including congestion control, routing, data analytics, mechanisms for low-latency responses, and application frameworks for IoT.

- **Usable Security and Privacy**: Bringing together computational scientists and social scientists, our research examines the confluence of human decision making and technology, such as wearables and IoT devices, their impact on privacy and security, and how to design systems that improve both privacy and security.

SOCIAL
email: info@icsi.berkeley.edu
web: www.icsi.berkeley.edu/icsi
twitter: @ICSIatBerkeley
linkedin:
https://www.linkedin.com/company/icsiatberkeley

SCIENTIFIC AND COMMERCIAL IMPACT

ISCI is built on a foundation of excellence and impact laid over the past three decades. In the early years, ICSI's research focused on speech, artificial intelligence, hardware, and theory, with each area producing seminal results.

ICSI research produced the first effective approach to handling spectral differences in audio channels used for speech recognition (RASTA); the first single-chip vector processor (Torrent); the first practical rateless codes and several successors; a novel neural theory of the basis of human language and thought; and a pioneering lexical database based on frame semantics (FrameNet).

More recent ICSI efforts in networking and systems, which involved collaborators at UC Berkeley and elsewhere, played a critical role in the development of: distributed hash tables, which were the forerunner of today's scalable key-value stores; software-defined networking, which is now widely used in commercial networks; network function virtualization, which moved critical network functionality into software; and Spark, a leading data analytics engine.

In some cases, the best way to bring these ideas to the world was through startups. Startups based on ICSI-related technology include Digital Fountain (founded in 1998, sold to Qualcomm in 2009), Nicira Networks (founded in 2007, sold to VMware in 2012), Xorp (founded in 2008), Databricks (founded in 2013), Corelight (founded in 2013), Nefeli Networks (founded in 2016), Appcensus (founded in 2019), and BitRipple (founded in 2020).

ICSI researchers have earned a wide range of awards which, in addition to numerous best paper and test-of-time awards, include the National Medal of Science, Turing Award, National Academy of Science membership (2), National Academy of Engineering membership (3), IEEE Hamming Medal, IEEE Sumner Award, ACM SIGCOMM Award (4), IEEE Internet Award (4), ACM Kanellakis Theory and Practice Award (2), and the IEEE Flanagan Speech and Audio Processing Award.

WHY JOIN ICSI?

The Institute is always looking for new researchers and collaborators—perhaps you! The Institute provides a flexible, open, and collaborative environment, uniquely suited to enable researchers to do their best work that can change the world.

- ICSI is in the heart of Berkeley, a block from the UC Berkeley campus, and within the Silicon Valley ecosystem. This provides easy access to students, collaborators, and entrepreneurial resources.
- ICSI has maintained a close affiliation with UC Berkeley since its founding in 1988. Several ICSI scientists hold joint faculty appointments at the university and many ICSI's scientists collaborate closely with research groups at UC Berkeley. These collaborations include: the UC Berkeley's Division of Computing, Data Science, and Society; the Berkeley Institute for Data Science; the EECS, Statistics, and Linguistics departments; and the I-School.
- Our researchers also collaborate widely with other universities, industry, government, and social impact organizations. Many also work with and have had joint appointments with government research laboratories, including Lawrence Berkeley National Laboratory, leading efforts in networking, machine learning, and other areas.
- Visitors travel from all over the world to pursue research at ICSI. The Institute has hosted scientists and more than 620 postdocs from Germany, Singapore, Spain, Switzerland, Finland, Italy, the United Kingdom, the Philippines, Israel, China, Japan, Brazil, Belgium, Canada, Holland, and more.
- We value—and don't hinder—interdisciplinary work that spans topics and even entire fields. In addition, ICSI encourages its researchers to pursue new and daunting challenges rather than "playing it safe". For instance, some of our recent efforts aim to bring about transformative changes in cloud computing, the Internet architecture, the cellular ecosystem, and Internet privacy.
- ICSI accommodates many different forms of appointments, from full-time researchers, to part-time researchers looking for an adjunct to their faculty, government lab or industry position, to visiting researchers who would like to collaborate with UC Berkeley or build connections in Silicon Valley.
- Unlike industry or academia, ICSI does not have strict expectations on research output. We may write traditional research papers, but we can just as easily release software, curate a dataset, or spin out a startup company.
- We value our employees' health and well-being. While located in downtown Berkeley, the Institute supports flexible, remote, and hybrid work options—at ICSI, you can choose to pursue research virtually from your home, the beach, or the mountains! We also offer generous health insurance and 401K benefits.

Donald Bren School of Information and Computer Sciences

UNIVERSITY OF CALIFORNIA, IRVINE

SUMMARY

As the only computational-focused school in the UC System, the UC Irvine Donald Bren School of Information and Computer Sciences has a unique perspective on the information technology disciplines that allows us a broad foundation from which to build educational programs and research initiatives.

PROGRAM STATISTICS

- Number of faculty: Split among departments
- Number of core staff: 3
- Number of affiliates: 50
- Number of students: 25

DEGREES, PROGRAMS, SPECIALIZATIONS

- Master of Data Science
- B.S. in Data Science

SOCIAL

email: mds@ics.uci.edu
web: mds.ics.uci.edu

LOCATION

Donald Bren School of Information and Computer Sciences
University of California, Irvine
6210 Donald Bren Hall
Irvine, CA 92697-3425

PROGRAM OVERVIEW

The Donald Bren's School of Information and Computer Science Master of Data Science (MDS) prepares you to develop your understanding of the cornerstones of modern data science through a dual framework in Statistics and Computer Science.

Our 15-month, full-time program taught by pioneering faculty and researchers in the field of data science anchors the curriculum on hands-on training in applied probability and mathematical statistics, statistical modeling and computing, machine learning, data management and visualization, and artificial intelligence.
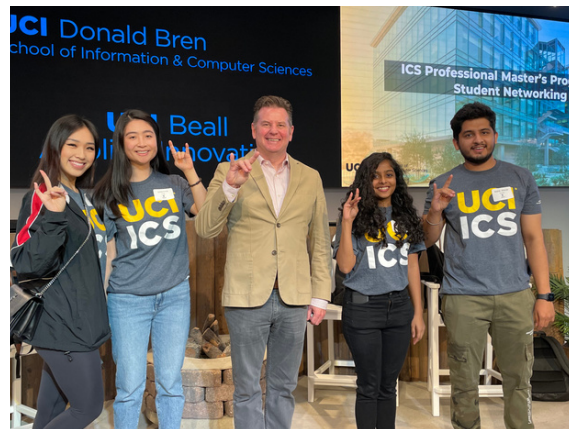
"Industry is learning quickly that people can scrape and analyze data, but to do something truly meaningful with data - to make decisions that will drive the industry forward you need the foundations of data science and understanding the statistics and computing methods. "

-Dan Gillen, Ph.D.
Chair of Department of Statistics

The MDS program provides an immersive education and experiential learning. Your time at the ICS School is focused on developing your skills and gaining exposure to the foundations of data science.

Between our Orange County corporate and tech connections and our combined faculty experience, the ICS School's global network is prominent. A large number of our alumni stay in Southern California, and our vast network spans across the world. No matter where your education takes you, the connections you make here will positively impact your life and career. We focus on the quality of professional and academic relationships that you build while here.

UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

SUMMARY

Founded in 2018 as an independent academic unit, the mission of the Halıcıoğlu Data Science Institute (HDSI) is to establish the scientific foundations of data science, develop new methods and infrastructure, and train students and partners to use data science to solve the world's most pressing problems.

PROGRAM STATISTICS

- Number of core faculty: 44
- Number of core staff: 15
- Number of affiliated members: 187
- Number of students:
  - 1450 undergraduate students in the major and minor
  - over 7000 enrolled in data science courses

DEGREES, PROGRAMS, SPECIALIZATIONS

- Doctor of Philosophy, Data Science
- Master of Science, Data Science
- Master of Data Science (Online)
- Bachelor of Science, Data Science
- Undergraduate Minor in Data Science

LOCATION

Halıcıoğlu Data Science Institute
University of California San Diego
10100 Hopkins Drive
La Jolla, CA 92093

TALENT PREPARATION IN DATA SCIENCE THROUGH EXPERIENTIAL LEARNING

HDSI's Data Science major preparation is designed to be transdisciplinary by providing a hub of instructional activities that draws faculty and researchers from different disciplines and departments. The program couples rigorous training in the theoretical foundations of data science with ample opportunities for experiential learning. It is enriched with real-world applications by partnering with industry and community organizations that provide access to rich data sets and varied use cases of societal importance while also increasing community engagement and facilitating multiple career pathways.

Experiential learning starts early with an Undergraduate Research Scholarship program and culminates in a senior capstone project designed to showcase mastery across the wide range of mathematical/statistical, technical, and conceptual skills developed throughout the curriculum. The capstone spans the entire data science life-cycle, including assessment of the problem, acquisition of domain knowledge, collection and cleaning of the data, system design, model creation and analysis, discussion of ethical implications, and presentation of results.

PARTNERING WITH THE SAN DIEGO SUPERCOMPUTER CENTER (SDSC) TO INCREASE UNDERSTANDING AND CAPACITY FOR SCALE AND COMPLEXITY

Founded as one of the original NSF Supercomputer Centers in 1985, SDSC remains at the forefront of high-performance computing. It provides the research community with large-scale compute and data resources, research networking, software, and computational and data science expertise.
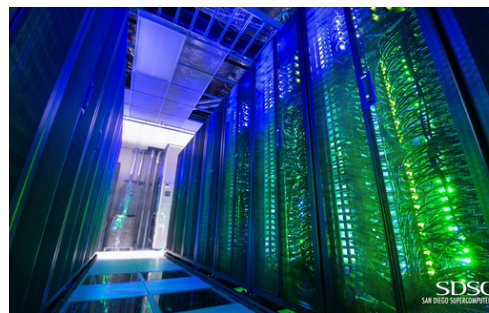
Today, SDSC's efforts are aimed at bringing together the power of machine learning, artificial intelligence, and data analysis. With simulation and large-scale experiments, SDSC supports the multidisciplinary approach needed to address critical research and societal challenges -- from astrophysics and Earth sciences to disease research and drug discovery. SDSC translates this innovation into practice through its collaborative partnerships across academia, industry, and the public sector.

HDSI and SDSC partner on programs for convergence research and translating data science innovations into practice.

CATALYZING LARGE SCALE INTEGRATIVE RESEARCH PROJECTS

HDSI provides intellectual and operational support that allow researchers drawn from various disciplines to work together in creating new research initiatives with potential for long-term impact. Among these efforts are the recently announced AI Institute, TILOS (The Institute for Learning Optimization at Scale), as well as DARPA efforts on Smart Radio platforms.

TILOS is a National Science Foundation funded National Artificial Intelligence (AI) Research Institute. The TILOS mission is to make impossible optimizations possible, at scale and in practice.
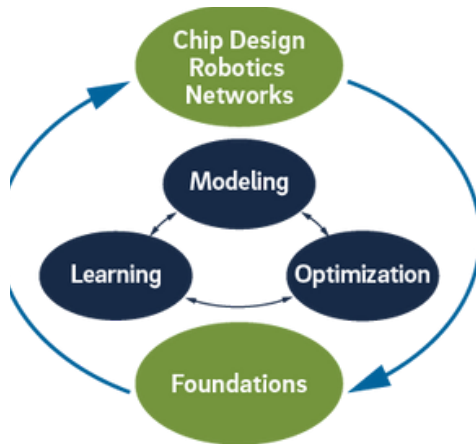
TILOS is a partnership of faculty from University of California, San Diego, Massachusetts Institute of Technology, National University, University of Pennsylvania, University of Texas at Austin, and Yale University. TILOS will pioneer learning-enabled optimizations that transform chip design, robotics, communication networks, and other use domains that are vital to our nation's health, prosperity and welfare. Foundational Research will pursue five main pillars:

1. Bridging discrete and continuous optimization.
2. Distributed, parallel, and federated optimization.
3. Optimization on manifolds.
4. Dynamic decisions under uncertainty.
5. Non-convex optimization in deep learning.

CREATING LEADERS IN DATA SCIENCE THROUGH GRADUATE STUDY

Three graduate programs join the Halıcıoğlu Data Science Institute's popular undergraduate Data Science programs for an integrated slate of courses and degree programs at all educational levels.

Master of Science: the MS program features courses in the foundational areas that prepare students from diverse backgrounds for a successful career in Data Science. A thesis option exists for students interested in cutting edge research.

Online Masters: the MDS program is designed for working professionals and combines concepts from statistics, computer science, and applications. The fully-online program is offered asynchronously in order to accommodate students with varying work schedules.

Doctor of Philosophy: the doctoral program creates leaders in the field of data science who will challenge and expand the boundaries of knowledge in the field. The doctoral program provides a research-oriented education spanning diverse data science areas, such as algorithms, machine learning, artificial intelligence, optimization, statistical methods, and data ethics.
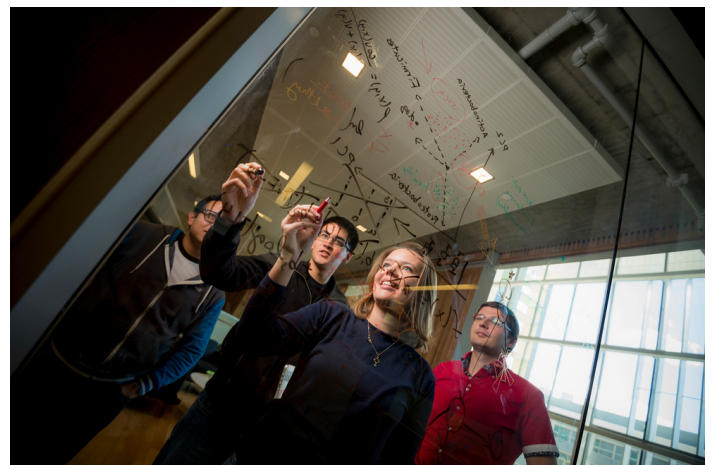
HDSI continues to develop innovative programs with other Departments, Schools, and Institutes.

Academic Data Science Alliance

CENTER FOR BLACK STUDIES RESEARCH

UC SANTA BARBARA

SUMMARY

We are an interdisciplinary vision of computing research that bridges Black Studies, Statistics, Molecular and Cellular Biology, Neuroscience, Environmental Science, Engineering and Computer Science to create a new and student-driven foundation through a data science living and learning community for understanding issues of personal significance through a diverse lens.

PROGRAM STATISTICS

- Number of core faculty: 2
- Number of core staff: 4
- Number of affiliated members: 17
- Number of students: 46

DEGREES, PROGRAMS, SPECIALIZATIONS

Students in this program are engaged in research across multiple departments. The data science program focuses on diverse epistemologies, translational and interdisciplinary along with an understanding of critical theory.

CBSR is in the process of getting certificates for this program

LOCATION

University of California, Santa Barbara; Center for Black Studies Research.
4603 South Hall
Santa Barbara, CA 93106

PROGRAM OVERVIEW

This data science program at UCSB is unique because it is housed in the Center for Black Studies Research and the focus is on various aspects of data science (visualizations, machine learning, neuro-computation, digital humanities, data analytics and statistics)using a critical theory framework. It is not a part of the formal curriculum. It combines, open discussions, workshops, research, research mentoring, presentations and critical readings.

UC SANTA BARBARA

UC SANTA BARBARA

SOCIAL

email: cbsr-assistant@ucsb.edu
web: cbsr.ucsb.edu/seeds

Academic Data Science Alliance

Center for Translational Data Science

UNIVERSITY OF CHICAGO

SUMMARY

The Center for Translational Data Science at the University of Chicago is developing the discipline of translational data science and applying it to tackle challenging problems in biology, medicine, healthcare and the environment.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Specializations in translational data science and its applications to biology, medicine, healthcare, and the environment.

PROGRAM STATISTICS

- Number of Faculty: 3
- Number of Staff: > 75 research scientists and professional staff
- Affiliated Faculty: 6

SOCIAL

email: ctds@uchicago.edu
web: ctds.uchicago.edu
twitter: @UChicagoCTDS
linkedin:
www.linkedin.com/company/center-for-translational-data-science
facebook:
www.facebook.com/UChicagoCTDS
youtube:
www.youtube.com/channel/UC_QDSvXzqQj6B5Vh6qiqT_Q

LOCATION

Center for Translational Data Science
University of Chicago
5454 South Shore Drive
Suite 2A/B
Chicago Illinois 60615

PROGRAM OVERVIEW

The Center for Translational Data Science has developed a number of important "firsts:" including, one of the first large scale data clouds (the NSF supported Open Science Data Cloud (2010-2016)); the first data cloud designed to host biomedical data and approved as a NIH Trusted Partner (the Bionimbus Protected Data Cloud (2013-present)); the first large scale data commons (the NCI Genomic Data Commons (2016-present)); and the first set of services to create data ecosystems for biomedical data (Data Commons Frameworks Services (2020-present)).

Today with our partners, we operate a data ecosystem compromising over a dozen data commons that make over 14 PB of data available to the research community on over 1,000,000 research participants. We provide access to this data via secure and compliant workspaces while protecting privacy.
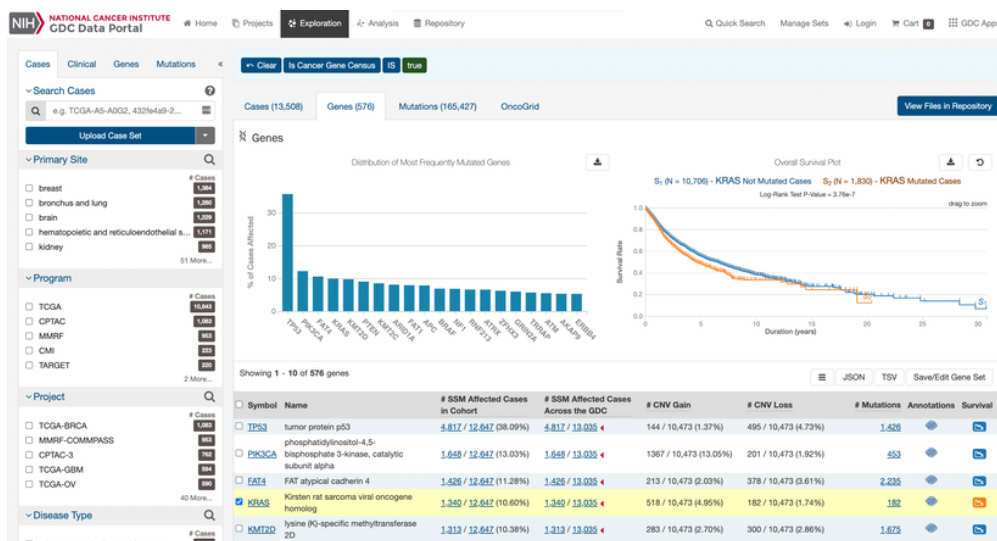
The NCI Genomic Data Commons, which was developed by CTDS

Academic
Data Science
Alliance

UNIVERSITY OF DELAWARE
DATA SCIENCE INSTITUTE

SUMMARY

The Institute aims to accelerate research in data science, serving as a nucleating effort to catalyze interdisciplinary research collaborations across fields impacting our society.

DEGREES, PROGRAMS, SPECIALIZATIONS

- MS in Data Science
- PhD/MS/Certificate in Bioinformatics Data Science
- MBA Major in Business Analytics
- PhD in Educational Statistics and Research Methods
- PhD in Financial Services Analytics (FSAN)
- PhD/MS/Certificate Geospatial Data Science
- And much more

PROGRAM STATISTICS

- Number of faculty: 102
- Number of staff: 13
- Number of Students: 162

SOCIAL

email: dsi-info@udel.edu
web: dsi.udel.edu
twitter: @dsiudel

LOCATION

Data Science Institute (DSI)
The Tower at STAR
100 Discovery Boulevard, Suite 614
Newark, DE 19713

INSTITUTE OVERVIEW

The University of Delaware's Data Science Institute (DSI) serves as a hub for **interdisciplinary research, collaboration, and excellence,** bringing together faculty and students from eight colleges across campus to work effectively with **big data** and address problems and opportunities facing society—from health sciences, physical sciences, engineering, environmental sciences to behavioral and social sciences and public policy.

INSTITUTE PURPOSE AND DESIGN

DSI aims to accelerate collaborative research in, and application of, data science. The DSI operates through four working groups (WGs): **Research WG** to foster multidisciplinary collaborations and joint grants, **Infrastructure WG** to provide data and compute infrastructure, **Training WG** to facilitate training and development opportunities, and **Networking & External Relations WG** to host events and connect faculty, students and industry and government partners.

DSI facilitates external access to the University's Data Science capabilities and assets, which include the **Data Intensive & Computational Sciences Core** and the **DARWIN High Performance Computing System**, and coordinates access to data-intensive research capabilities including bioimaging and other Core research facilities at the University. The institute further involves partnerships with industry and other institutions in the region.

OUR FACULTY

DSI brings together more than 100 faculty from all eight colleges and schools across campus to work collaboratively in a wide range of topics in foundations and applications of data science. DSI faculty at UD combine expertise in statistics, computer science, mathematics, information sciences and numerous related fields.
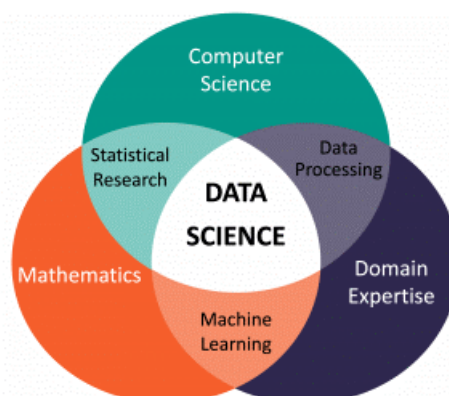
13 Resident Faculty across the colleges were hired as a Presidential Strategic Initiative to complement the **100+ DSI Affiliated Faculty**. Resident Faculty lead the Certificate in Urban Data Science, PhD in Educational Statistics and Research Methods, and the NSF-funded Delaware and Mid-Atlantic Data Science Corps.

DATA SCIENCE TRAINING AND EDUCATION

The interdisciplinary data science programs are offered through 11 graduate, undergraduate, and certificate programs including **76 data science-related courses**.

With a flexible set of core and elective courses, the data science degrees lead to a wide range of potential application areas. By providing a solid background in the methods behind data science, the programs enable our graduates to work well with data and be better prepared for the latest methods of focusing on large or dynamic data sets in their chosen fields.

Our degree programs and training grants such as **NSF NRT-HDR & HDR-DSC grants** are preparing students for an extensive variety of data science positions. Data analysts use mathematical, statistical and modeling techniques to solve problems; data engineers design, build and maintain an organization's data and analytical infrastructure; and data scientists create sophisticated analytical models to build datasets and derive new insights from data.

INDUSTRY ENGAGEMENT

DSI is establishing an **Industry Affiliate Program** to broaden industry engagement. Leveraging expert faculty across the University and cutting-edge data science capabilities and unique resources, industrial partners can:
- Address industry-wide problems through collaborative research projects
- Demonstrate, test, & improve high-risk / high-return concepts in a protected environment
- Gain access to powerful computational resources & expertise
- Participate in & co-host special events, e.g., lectures, seminar series, workshops, trainings, hackathons
- Access resources for workforce development, including certificates, short courses, and degree programs

DATA-INTENSIVE & COMPUTATIONAL SCIENCE (DICOS) CORE FACILITY

The **Data Intensive & Computational Science (DiCoS) Core** supports interdisciplinary research collaborations and team science. DiCoS aims to advance both the research frontiers and the underlying infrastructure at the nexus of computational and data science, enabled by high-performance computing and big data.

Computing Hardware and Software Resources

Storage, Archiving and Curation

Research Data Sets and Software Tools

Data Science and Software Engineering Expertise

The Data Intensive & Computational Science (DiCoS) Core supports interdisciplinary research collaborations and team science. DiCoS aims to advance both the research frontiers and the underlying **infrastructure at the nexus of computational and data science**, enabled by high-performance computing and big data.

Academic Data Science Alliance

National Center for Supercomputing Applications

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

SUMMARY

The National Center for Supercomputing Applications (NCSA) at the University of Illinois Urbana-Champaign is a campus-wide interdisciplinary center. Since 1986, we've been at the epicenter of supercomputing research, pioneering innovations in computing and data. Our advanced cyberinfrastructure and expertise provide a hub for transdisciplinary research for both academia and industry.

PROGRAM STATISTICS

- Number of core professional staff: 240
- Affiliated researchers: 131

SOCIAL

email: info@ncsa.illinois.edu
web: www.ncsa.illinois.edu
twitter: @NCSAatIllinois
linkedin: www.linkedin.com/company/ncsaatillinois
facebook: www.facebook.com/NCSAatIllinois
instagram: @ncsaatillinois
youtube: NCSAatIllinois

LOCATION

Urbana, Illinois USA

PROGRAM OVERVIEW

The National Center for Supercomputing Applications (NCSA) is one of 10 interdisciplinary campus institutes at the University of Illinois Urbana-Champaign. For more than 35 years, NCSA has been applying advanced computing to change the world. NCSA currently is home to multiple programs in data science, including our Center for AI Innovation, Advanced Visualization Laboratory, Center for Astrophysical Surveys, and Health Innovation program office. NCSA's Industry program offers data science expertise to the private sector. NCSA also hosts programs that provide opportunities and training for students in data science, an REU site and DIGI-Mat, a National Science Foundation Research Traineeship program in cooperation with the department of materials science. NCSA also hosts the Midwest Big Data Innovation Hub (MBDH), one of four NSF regional big data hubs. The focus of NCSA has always been on applications, and today many applications build on combining data and computing. NCSA also operates powerful supercomputers, including the newly installed Delta, optimized for machine learning as well as simulations.

MIDWEST BIG DATA INNOVATION HUB

The MBDH is based at NCSA and convenes and coordinates data science communities across the region through a variety of activities. including:

- Advanced Materials and Manufacturing – building capacity by developing effective pathways to incorporate data into the materials and manufacturing space, such as integrating existing materials and manufacturing-related software and services
- Digital Agriculture – making the data that describe ecosystems, crops, and animals more easily findable, accessible, interoperable, and reusable (FAIR) to improve the agriculture process and its goods.
- Smart & Resilient Communities – shaping conversations on equitable data use, innovative research, and community-driven, data-informed planning and decision-making to address the challenges of using infrastructure, safety and security, and coexistence of diverse communities

- Water Quality – building and convening an innovation network to accelerate and sustain the development of standards, data sets, and cyberinformatics tools to address critically missing links and bottlenecks in water-resources research and development
- Big Data in Health – developing infrastructure enabling powerful and secure mechanisms for data aggregation and open sharing to manage, harmonize, aggregate, model, interrogate and visualize complex biomedical and health data

CENTER FOR ARTIFICIAL INTELLIGENCE INNOVATION

NCSA's Center for Artificial Intelligence Innovation leverages existing partnerships and forges new collaborations to solve difficult multidisciplinary challenges and expand awareness of what AI can achieve. Maximizing synergies between industry teams, students, computing experts, and researchers enables powerful, adaptive, lasting solutions to intractable problems.

CENTER FOR ASTROPHYSICAL SURVEYS

A collaborative effort between NCSA, and the university's Office of the Vice Chancellor for Research and Innovation, department of astronomy, and department of physics, CAPS projects include supporting the Dark Energy Survey, helping researchers explore the accelerating expansion of the universe; the Vera C. Rubin Observatory working to discover the structure and evolution of the universe; and support of the South Pole Telescope and others around the world.

ACADEMIC PROGRAMS

The University of Illinois Urbana-Champaign has many strong academic programs in data science and related disciplines. For example, the department of computer science offers an online Master of Science in Data Science; the department of civil and environmental engineering offers an MS with a data science track; statistics offers a Certificate in Data Science; among many other such programs on campus. Illinois is also home to many successful and established degree programs in "CS+X", where X ranges from mathematics to music. Also notable are the programs in the School of Information Sciences, which has led to partnerships with NCSA such as the Community Data Clinic, and which includes leading a multi-departmental MS in Bioinformatics, and both a Minor and PhD in Informatics. Focusing on industry, the Gies College of Business offers an MS in Business Analytics. Illinois also has strong programs in digital humanities, both through the Institute for Research in the Humanities (another campus-wide interdisciplinary institute) and through academic programs in numerous departments.

DEPARTMENT OF INFORMATION SYSTEMS

SUMMARY

UMBC is a dynamic, R1, public research university integrating teaching, research and service to benefit the citizens of Maryland and the world. Our UMBC community redefines excellence in higher education through an inclusive culture that connects innovative teaching and learning, research across disciplines, and civic engagement.

PROGRAM STATISTICS

- Number of faculty: 35
- Number of core staff: 5
- Number of students: 1600+ total students enrolled in our graduate and undergraduate programs

DEGREES, PROGRAMS, SPECIALIZATIONS

- Master of Science in Information Systems - On Campus
- Master of Science in Information Systems - Online
- Bachelor of Science in Information Systems
- Bachelor of Arts in Business Technology Administration
- Certificates in:
  - Cybersecurity Informatics - Undergraduate
  - Decision Making Support - Undergraduate

LOCATION

1000 Hilltop Circle
ITE 404
Baltimore, MD 21250

PROGRAM OVERVIEW

The Department of Information Systems is part of The College of Engineering and Information Technology (COEIT) at UMBC. We currently offer a variety of degrees and certificates for undergraduate and graduate students, including an U.S. News ranked Online M.S. in Information Systems and specializations in data science.

Through our curriculum, Information Systems students investigate societal impact, keeping humans in the loop, dealing with real and imminent challenges facing society. With the help of our faculty members, students study, design, develop, and evaluate information technologies to address the needs of a broad range of individuals and organizations.

DATA SCIENCE SPECIALIZATIONS

Businesses and governments around the world are using big data to make big decisions, which is why we have integrated data science throughout our programs, including specialized track offerings. Our Online and On-Campus Master of Science in Information Systems degrees both include tracks in Artificial Intelligence and Data Science. These tracks appear of a student's transcript and help give them an advantage when seeking employment or terminal degrees.

Our AI track includes courses like Deep Learning and Social Media Application and Analysis. It teaches students how to make machines find the data they want, derive meaningful information from it, and teach it to make decisions based on that data. Our Data Science track includes courses like Information Integration and Data Analytics in Cybersecurity. This track teaches students the fundamentals of data science and how to apply it in different contexts. These skills are vitally important in the 21st century, across all industries.

DEPARTMENT OF INFORMATION SYSTEMS

CUTTING-EDGE RESEARCH

UMBC is a top-tier, R1, public research university leading the world in inclusive excellence in research. Our research aims to further innovation in Information Systems through inclusive interdisciplinary partnerships and collaborations. We accomplish this by recruiting and retaining faculty with a passion for civically engaged research.

We are committed to mentoring students in research. Participating in research helps students develop their critical thinking, build transferable skills, and explore potential careers. The mentoring relationships built between students and faculty can last a lifetime and benefit both parties. We get students involved as early as and often as possible in research

Our main research areas in the IS department are AI and Machine Learning, Data Science, Health Informatics, Human-Computer Interaction, and Software Engineering.
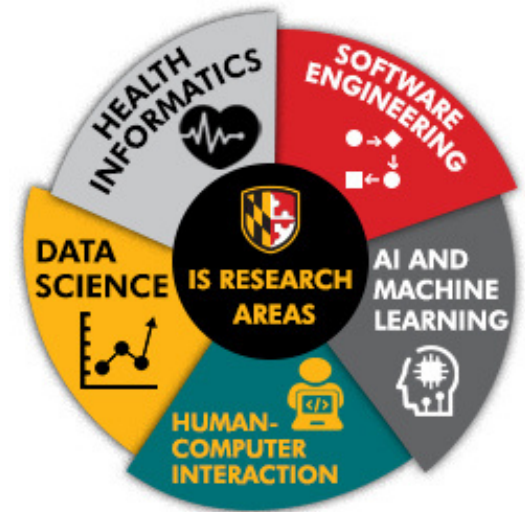
DATA SCIENCE MEETS CLIMATE SCIENCE

Tens of millions of people live in areas that are at risk for flooding due to climate change, sea level rise, and melting of glaciers. The Department of Information Systems, along with a team of researchers, are using data science, machine learning, and artificial intelligence (AI) to analyze enormous volumes of climate data, and Arctic and Antarctic observations in ways that could help populations prepare for and respond to these risks. Their research is part of a new five-year, $13 million grant from the National Science Foundation's Harnessing the Data Revolution (HDR) Big Idea program.

Center for Data Science

UNIVERSITY OF MASSACHUSETTS AMHERST

SUMMARY

The Center for Data Science fosters research, education, industry collaboration, and public service to make UMass Amherst a destination and partner-of-choice for research in data science. We help companies and organizations meet their growing demand for well-trained data scientists, promote economic development, and support working data scientists in Western Massachusetts.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Masters Concentration in Data Science
- Certificate in Statistical and Computational Data Science

PROGRAM STATISTICS

- Number of faculty: 43
- Number of staff: 6
- Number of Students: approximately 345

SOCIAL

email: cds-info@cs.umass.edu
web: ds.cs.umass.edu
twitter: @UMassDataSci

LOCATION

Center for Data Science
University of Massachusetts, Amherst
740 North Pleasant Street, Suite A205
Amherst MA 01003

DATA SCIENCE FOR THE COMMON GOOD

CDS is leading efforts to provide education and research pathways for aspiring data scientists to apply their knowledge and skills to benefit society. Data Science for the Common Good™ (DS4CG) is a summer program that trains aspiring data scientists to work on real-world problems that benefit the common good. Our teams of computer science Master's students collaborate with nonprofit organizations and government agencies working in public health, education, health and wellness, environmental conservation, and more. DS4CG harnesses growing student interest in social-good causes and connects it to partner organizations that stand to benefit from the students' growing data science expertise.

DS4CG serves organizations that lack staff capacity to exploit state-of-the-art analysis and modeling of their business data. Ideal DS4CG projects incorporate analysis and integration of very large volumes of data, with different types and formats. DS4CG fellows compile, organize and clean these datasets, mathematically explore their statistical properties, build predictive models and forecasts, and visualize the results. Project deliverables can include actionable data-driven insights, presentations, reports, and proof-of-concept software tools. Learn more here: https://ds.cs.umass.edu/industry/data-science-common-good

EQUITY, ACCOUNTABILITY, TRANSPARENCY, AND EXPLAINABILITY

EQUATE is an acronym for Equity, Accountability, Transparency, and Explainability. These areas are commonly referred to as FAT (fairness, accountability, and transparency).

EQUATE is an interdisciplinary initiative of UMass Amherst faculty who are engaged in research and education related to equitable algorithms and systems. The initiative is supported by the Center for Data Science, and many of the EQUATE community members are drawn from CDS-affiliated faculty in the College of Information and Computer Sciences. The group's educational efforts include coursework in ethics and algorithm design that respects the values of fairness and transparency. Community research efforts explore EQUATE topics within software systems and programming languages, machine learning, and vision, theory, and data management systems.

Learn more here: https://ds.cs.umass.edu/research/equate-equity-accountability-transparency-and-explainability

Center for Data Science

UNIVERSITY OF MASSACHUSETTS AMHERST

DATA SCIENCE INDUSTRY MENTORSHIP PROGRAM

The Data Science Industry Mentorship Program is an exclusive benefit of the Center for Data Science Industry Affiliates Program. The program matches small teams of data science Master's students with an industry-proposed project. Over the course of an academic semester each team works under the guidance of an industry mentor. Program Objectives:

* Small teams of MS-level data science students at UMass Amherst get the opportunity to work on industry-relevant problems, with guidance and mentoring from industry data science professionals.

* Companies get the opportunity to make cost-effective progress on data science exploratory problems of interest, leveraging the effort of students who are in the midst of data science training. Company professionals "learn alongside" the student teams.

* As a result of the experience of working with these students, participating companies may find candidates for future internships and full-time roles.

Past runs of this course have resulted in publications at high-impact conferences, and some have received best papers awards.

Learn more here: https://ds.cs.umass.edu/industry/industry-mentorship-program

DATA SCIENCE AND SOFTWARE ENGINEERING CORE

The Data Science and Software Engineering Core provides data science and related software engineering services in support of research and development. Services offered include analysis, application development, deployment, and prototyping in areas such as data science, security, forensics, software development, informatics, machine learning, computer vision, natural language processing, and cloud computing. Building on the foundation of the world class expertise of computer science faculty, we provide a deep and dedicated team of professional scientists and engineers that can be shared widely, to solve a range of research, application, and cross-disciplinary problems. We foster engagements at all levels from supporting basic research, to solving applied problems, to implementing models, software, and solutions at scale. Services are provided to the Manning College of Information and Computer Sciences, to the University as a whole, as well as to outside government, non-profit, and for-profit enterprises.

Michigan Institute for Data Science

UNIVERSITY OF MICHIGAN

MIDAS

SUMMARY

MIDAS strengthens University of Michigan's research capacity in Data Science and Artificial Intelligence, and enables their transformative use for scientific discovery and lasting societal impact. Its faculty community includes 420 methodologists and domain scientists from all schools and colleges at the Ann Arbor campus, and Dearborn and Flint campuses.

PROGRAM STATISTICS

- Number of technical staff: 2.5
- Affiliated Faculty: 430
- Students: 73 currently enrolled in the certificate program

DEGREES, PROGRAMS, SPECIALIZATIONS

- Graduate Data Science Certificate

SOCIAL

email: midas-contact@umich.edu
web: midas.umich.edu
twitter: @um_midas
linkedin: www.linkedin.com/company/michigan-institute-for-data-science-midas

LOCATION

500 Church Street, Suite 600
Ann Arbor, MI 48109-1042

ETHICAL DATA SCIENCE AND AI

As data science and AI become a major force in science and in society, increasingly complex analytical pipelines working with poorly understood data pose significant issues of bias, inclusion and fairness. MIDAS is mobilizing our researchers to promote ethical data science and AI. The highlights include:

Raising awareness through seminars, community forums and public exhibits. Speakers such as Timnit Gebru (founder of Black in AI), Catherine D'ignazio and Lauren Klein (authors of Data Feminism) sparked campus-wide discussions on how to correct data bias. A MIDAS-led project, "White Cube, Black Box", currently on exhibit at the University's Museum of Art, illustrates biases in face detection algorithms and in museum practices of artwork acquisition.

Technical solutions. MIDAS Director, H.V. Jagadish, leads a multi-university team to develop a system that will take potentially biased data from both public and private sources, and make them bias-adjusted and analysis-ready. The project has completed initial work on methods to identify data and software bias that negatively impact equity in critical domains, including mobility, housing, education, economic indicators and government transparency.

REPRODUCIBLE RESEARCH

Amid the worldwide outcry to address the "reproducibility crisis" in scientific research, funding agencies, journals, professional societies and individual researchers are all stepping up to offer solutions and incentives. Data science institutes have a unique role to play to enable reproducibility. MIDAS highlights include:

Reproducibility Challenges. We organized this unique series in 2020 and 2021, in which researchers submitted their approaches to improving the reproducibility of data-intensive research. These included methodological developments for more robust (hence more reproducible) study designs, tools and workflows to reproduce projects, reproducing published works, meta-analysis of reproducibility issues in a research field, and teaching materials.

Reproducibility Resources. Two key observations emerged from the Reproducibility Challenges: 1) many researchers develop their own tools, but don't have the resources to improve, maintain and standardize them; and 2) many researchers care about the reproducibility of their work but are unsure of what tools to use. MIDAS has put together a Reproducibility Showcase series and an online resources collection. We believe that providing practical solutions is a niche that data science institutes should fill.

TRANSFORMING DISCIPLINARY AND ENABLING INTERDISCIPLINARY RESEARCH

As a campus-wide organization, MIDAS enables the transformative use of data science and AI methods in traditional disciplines and facilitates interdisciplinary research through catalyzing ideas and teams. The highlights include:

Seed funding for groundbreaking research. MIDAS funded 52 projects in the past six years. The funding enabled ~90 external grants with the total funding amount 9 times the amount of MIDAS funding. The seed funding supports methodology development as well as the application of cutting-edge data science and AI methods to a wide range of research domains, including biomedical and healthcare research, engineering, environmental and sustainability, learning analytics, and social sciences. Interdisciplinary teams and approaches are particularly favored for seed funding.

Convening researchers for convergence science. Taking advantage of the scientific diversity of our community, MIDAS organizes research working groups to address data and analytics issues common to many disciplines, with themes such as "sensor data analytics", "missing data in COVID research" and "biology inspired math theory." Through working groups and seed funding, MIDAS also brings together diverse perspectives to address significant scientific and societal challenges, for example, convening environmental scientists, social scientists, data science methodologists and medical researchers to study pandemic prevention. Such activities have proven to be a fertile ground for highly interdisciplinary research ideas and teams.

COLLABORATION WITH ACADEMIA, INDUSTRY AND THE PUBLIC SECTOR

MIDAS is also the university's "front door" for external collaborators of data science and AI, and has built extensive collaboration with academic data science institutes, industry, government and community organizations to advance research and promote Data for Social Good. The highlights include:

Future Leaders Summit to promote collaboration among data science institutes and to foster the careers of the future research leaders. Participants of this annual event are PhD students and postdocs from universities with mature data science programs as well as those beginning to build such programs, from major research universities to minority serving institutions. The event includes research presentations, career mentoring and networking sessions. More than half of all participants are women and underrepresented minorities. The theme of the 2022 Summit is "Responsible Data Science and AI."

Data for Social Good. MIDAS researchers and students carry out data science projects to support the data strategy of government and community partners, including multiple organizations in the City of Detroit and the City government, the Native American tribal nations in Michigan and other organizations in Michigan. These projects also allow researchers to apply ethical data science approaches for positive societal impact.

UNC-Chapel Hill/Renaissance Computing Institute (RENCI)

SUMMARY

RENCI is a highly collaborative research institute at UNC-Chapel Hill that develops and deploys advanced technologies to enable innovative research and discovery in high performance computing (HPC), informatics, data mining, data linking, and secure computing.

PROGRAM STATISTICS

- Number of core staff: 51 research and teaching; 94 total employees
- Number of affiliated members: 5

DEGREES, PROGRAMS, SPECIALIZATIONS

- Specializations in:
  - Machine Learning
  - Biomedical Science
  - Health Informatics
  - Earth Data Science
  - Networking & Security
  - Data Mining & Linkage

SOCIAL

email: comms@renci.org
web: renci.org
twitter: @RENCI
linkedin:
www.linkedin.com/company/renaissance-computing-institute
facebook:
www.facebook.com/renci.org
youtube:
www.youtube.com/channel/UCSOkatGqnWS_o1rU1mYgxrA

LOCATION

100 Europa Dr, Suite 540
Chapel Hill, NC 27517

RENCI'S DATA SCIENCE COMPUTATIONAL PLATFORM

RENCI researchers have developed a flexible computational workspace for any scientific domain. The tool allows researchers to bring together tools specific to their work in a secure, scalable portal. A version tailored for educational use, EduHeLx, has been developed and was deployed in a pilot capacity in Fall 2021 in the UNC-Chapel Hill course, *COMP 116: Introduction to Scientific Programming*. EduHeLx is now being used again in the same course in Spring 2022. As UNC-Chapel Hill continues to develop its data science school, EduHeLx has the potential to serve as a primary data science computational platform for data science education at the University.

SOUTH BIG DATA HUB

The South Big Data Hub (South Hub or SBDH) serves the Southern U.S. Census Region (16 States: Delaware - Texas, including Washington DC, Puerto Rico, the U.S. Virgin Islands, and territories). It is part of a network of four regional Big Data Hubs, launched by the National Science Foundation, including the Midwest, Northeast, and West Big Data Hubs, to engage local or regional stakeholders in big data research and permit a focus on regional issues. The South Hub builds the necessary bridges and acts as a community convener, collaboration hub, and catalyst for data science innovation, through education, funding support, and collective action. The South Hub also amplifies the successes of the community, provides programs and funding, as well as shares credit across the community to encourage collaboration and mutual success toward shared Goals and Priority Areas. The South Hub is managed jointly by the Georgia Institute of Technology and RENCI, with more than 1300+ members from universities, corporations, foundations, and cities. For more information, visit southbigdatahub.org/about.

NHLBI BIODATA CATALYST

NHBLI BioData Catalyst (BDCatalyst) is a cloud-based platform providing tools, applications, and workflows in secure workspaces. By increasing access to NHLBI datasets and innovative data analysis capabilities, BDCatalyst accelerates efficient biomedical research that drives discovery and scientific advancement, leading to novel diagnostic tools, therapeutics, and prevention strategies for heart, lung, blood, and sleep disorders.

RENCI serves as the Coordinating Center for BDCatalyst in partnership with RTI International. The Coordinating Center functions to manage and unify the widespread and diverse teams working to build the ecosystem. For more information, visit https://biodatacatalyst.nhlbi.gov/.

NATIONAL CONSORTIUM FOR DATA SCIENCE (NCDS)

The National Consortium for Data Science (NCDS), founded by RENCI, is a collaboration of leaders in academia, industry, and government formed to address the data challenges and opportunities of the 21st century. The NCDS helps members take advantage of data in ways that result in new jobs and transformative discoveries. We connect diverse communities of data science experts to support a 21st century data-driven economy by: 1) Building data science career pathways and creating a data-literate workforce, 2) Bridging the gap between data scientists in the public and private sectors, and 3) Supporting open and democratized data. For more information, visit datascienceconsortium.org.

IRODS CONSORTIUM

The iRODS Consortium, founded by RENCI, brings together businesses, research organizations, universities, and government agencies to ensure the sustainability of iRODS by:

- Guiding further development of the software;
- Growing the user and developer communities; and
- Facilitating iRODS support, education, and collaboration opportunities.

The iRODS Consortium maintains and supports a commercial-grade distribution of iRODS. Additionally, the Consortium fields a team of software developers, application engineers, and support staff housed at RENCI. Each year, the Consortium hosts the iRODS User Group Meeting, a symposium that draws 100+ participants to share iRODS technologies and case studies. For more information, visit irods.org.

NCATS BIOMEDICAL DATA TRANSLATOR

The NCATS Biomedical Data Translator project applies semantic integration strategies to share chemical, genetic, phenotypic, disease, ontological, and other knowledge sources. Taken together, they form a technology platform for translational science. RENCI contributes to three Translator Projects: the Exposures Provider, Ranking Agent, and Standards and Reference Implementations (SRI) teams. The Exposures Provider team has developed ICEES (Integrated Clinical and Environmental Exposures Service) as a regulatory-compliant framework and approach for openly exposing and sharing integrated clinical and environmental exposures data. ICEES was designed as a general approach to overcome the numerous regulatory, cultural, and technical challenges that hinder efforts to openly share clinical data or any data containing personally identifiable information (PII). The Ranking Agent team has produced a tool called ARAGORN, which represents the next iteration of its successful ROBOKOP tool. ARAGORN answers user-specified biomedical questions by reasoning over a federated knowledge-graph, applying on-the-fly ranking to produce the most accurate results quickly. Additionally, ARAGORN attempts to bridge the precision mismatch between data representations and algorithms that require specificity, and users who pose questions and prefer answers at a more abstract level. The Standards and Reference Implementations (SRI) team produces a suite of standards and products, a model for their governance, and processes to coordinate integration and shared implementation

SUMMARY

The School of Data Science at UNC Charlotte commits to excellence in education, research, community engagement, and inclusion to shape and lead the future of data science education. We teach students to be responsible and ethical data science practitioners, leaders, and researchers in an increasingly data-driven and global society.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Bachelor of Science in Data Science
- Data Science and Business Analytics Master's Degree
- Health Informatics and Analytics Master's Degree
- Sports Analytics Certificate
- Data Science and Business Analytics Graduate Certificate
- Health Informatics and Analytics Graduate Certificate
- Specializations in Engineering, Health and Human Services, Computer Science, Business, Humanities, Social Science, Statistics

PROGRAM STATISTICS

- Number of Faculty: 14
- Number of Staff: 6
- Affiliated Faculty: 69
- Students:
  - 240 - DSBA graduate students
  - 96 - Data Science Undergraduate Students
  - 72 - Health Informatics and Analytics Graduate Students

LOCATION

Bioinformatics Building
9331 Robert D. Snyder Road
Charlotte NC 28223

PROGRAM OVERVIEW

Our BS in Data Science is fully integrated with the social sciences. The undergraduate curriculum is built around four interdisciplinary studio classes. Studio courses are team-taught by one faculty recruited from a technical background and a second from a specific content domain. Faculty from Software and Information System and Computer Science are partnered with Faculty from Sociology, Ethics, Political Science, and Criminal Justice to deliver the lab-based studios. Students work in teams to apply the tools of data science to real-world, community-based projects. The complexity of the problems increases at each level culminating within the fourth-year capstone studio, Data Science for Social Good, where students combine technical, analytic, interpretive, and social dimensions to design and execute a full data science project.

ETHICS IN DATA SCIENCE

Ethics was designed into data science courses and scaffolded learning objectives across all four years. As an interdisciplinary program rooted in a five-college partnership, our commitment is to incorporate liberal arts and science principles into a traditional math and computing-heavy major, specifically to address complex ethical issues embedded

ETHICS IN THE DATA SCIENCE CLASSROOM
SDS faculty understand the harm data science can cause and are teaching their students to do the same.

in data science. "One hundred percent of our data science courses at the undergraduate level address the ethical issues in data science," said Dr. Angela Berardinelli, senior lecturer for the College of Computing and Informatics and the School of Data Science.

"Other programs tend to offer a single elective on ethics or create a module within an introductory or advanced course." The major not only informs students about data ethics in the classroom, but will also enable them to engage with community partners to observe the impact of data, privacy, and their work in real life. Students begin with the basics of data ethics, gradually adding the evaluation of ethical debates and arguments, then concluding with learning how to conduct an ethical audit of real-world scenarios within data science.

INDUSTRY PARTNERSHIPS

We have very strong industry partnerships. Community is one of the four central pillars in the School of Data Science's (SDS) mission statement. Specifically, our goal is to, "Engage the Charlotte and campus communities by creating accessible pathways for research and practice." From the inception of our first degree offering in Health Informatics, we have relied on the support and counsel of an industry advisory board made of industry and non-profit leaders in the region. As we added degree programs and the Data Science Initiative has grown into the School of Data Science, our Industry Advisory Board has evolved. Over 25 companies and community organizations are represented on our board. They are local, regional and national entities that represent healthcare, financial services, energy, manufacturing, and retail. Our success is predicated on their support and engagement.

INTERDISCIPLINARY PROGRAM

We are structured as an interdisciplinary School that is jointly governed by 5 colleges. Within the academic sphere, the School of Data Science is governed by an Academic Advisory Board made up of five colleges: The Belk College of Business, The College of Computing and Informatics, The College of Health and Human Services, and the William States Lee College of Engineering. The SDS Faculty is composed of more than eighty Affiliate, Joint, and Core faculty.

This interdisciplinary collaboration reaches into almost every college and a majority of academic departments at UNC Charlotte.

SOCIAL

email: datascience@uncc.edu
web: datascience.charlotte.edu
twitter: @CLTDataScience
linkedin:
www.linkedin.com/school/unc-charlotte-school-of-data-science
facebook:
www.facebook.com/UNCCharlotteSDS
instagram: @cltdatascience

![OU DATA INSTITUTE FOR SOCIETAL CHALLENGES - The UNIVERSITY of OKLAHOMA](logo)

SUMMARY

Data Science Institute for Societal Challenges (DISC) is creating innovations in data science, artificial intelligence (AI), machine learning (ML), and data-enabled research. DISC develops and grows convergent research teams dedicated to solving local to global-scale challenges.

PROGRAM STATISTICS

- Number of directory members: 265
- Number of affiliated members: 88

DEGREES AND PROGRAMS AT OU

- OU offers a variety of undergraduate and graduate majors, minors, and certificates in data science and other related fields.

DISC SPECIALIZATIONS

- Data Science Applications
- Human-Guided Artificial Intelligence and Machine Learning
- Human-Computer Teaming
- Predictive Analytics
- Decision-Making Environments and Visual Analytics
- Scalable Software and Hardware Architectures

LOCATION

DISC
Five Partners Place
201 Stephenson Pkwy, Ste 4600
Norman, OK. 73019

DISC'S ROLE AT UNIVERSITY OF OKLAHOMA

DISC is playing a key role in OU's strategic plan. The Institute is integrated into and provides foundational capabilities to enable advancing the four research strategic themes: Aerospace, Defense and Global Security; Environment, Energy, and Sustainability; the Future of Health; and Society and Community Transformation. The four research strategic themes were created to address specific global challenges with convergent research partnerships using OU's expertise. We believe that this along with the integration of DISC with the University's strategic plan, ensures focus, relevance, and opportunities for both impact and success.

DISC COMMITMENT TO DIVERSITY, EQUITY, AND INCLUSION

The Data Institute for Societal Challenges (DISC) is committed to achieving a diverse, equitable, and inclusive data science community by embracing and valuing each person's unique contributions, background, and perspectives. We acknowledge that data science and data-enabled research have been used to reinforce unjust social, political, and economic systems that have resulted in systemic discrimination and marginalization of certain groups. DISC willactively work to ensure that the scholarship in which we engage does not perpetuate these inequities and instead seeks to reduce them. We will develop inclusive, trust-based, collaborative teams that empower the voices of underrepresented groups. We will identify and help address the issue of inequities and sources of bias in the broader data science community. Our work on developing convergent solutions to societal challenges includes a focus on equity.

SEED FUNDING TO ADVANCE RESEARCH AT OU

DISC created a research seed funding program designed to incentivize transdisciplinary, convergent research teams focused on tackling grand challenges. Seed funding allows OU researchers to incubate ideas with the potential for future extramural support. To date, DISC has awarded over $180,000 in seed funding.

SOCIAL

email: disc@ou.edu
web: ou.edu/disc
facebook: www.facebook.com/DISCatOU

Penn Arts & Sciences
Data Driven Discovery

SUMMARY

Penn's School of Arts and Sciences launched the Data Driven Discovery Initiative in 2021 to promote the development and use of data science science in the physical sciences, life sciences, social sciences, and humanities.

PROGRAM STATISTICS

- Number of core staff: 1
- Number of affiliated faculty: 29

DEGREES, PROGRAMS, SPECIALIZATIONS

- PhD in Computing & Data Sciences
- BS in Data Science
- Undergraduate Minor in Data Science
- All programs feature methodological and in-the-field tracks

SOCIAL

email: ddd-info@sas.upenn.edu
web: web.sas.upenn.edu/data-science

LOCATION

McNeil Building 558
3718 Locust Walk
University of Pennsylvania
Philadelphia, PA 19104-6286

PROGRAM OVERVIEW

At this initial stage we are launching a post-doc program, data science for social good seed grants, a data science minor, workshops and tutorials, and summer programming for Arts & Sciences undergraduates.

GOERGEN INSTITUTE FOR DATA SCIENCE

UNIVERSITY OF ROCHESTER

SUMMARY

Established in 2015, the Goergen Institute for Data Science (GIDS) serves as University of Rochester's integrative data science hub. The Institute offers a variety of data science degree programs, supports interdisciplinary data science research, and fosters industry-academia data science collaborations.

PROGRAM STATISTICS

- Number of core professional staff: 4
- Affiliated members: 79
- Total Enrolled Students: 263 - includes 183 BA/BS students, 67 MS students, and 13 advanced certificate students

DEGREES, PROGRAMS, SPECIALIZATIONS

- Bachelor of Arts (BA)
- Bachelor of Science (BS)
- Master of Science (MS)
- Advanced Certificate
- BA/BS specializations: biology, biomedical signals and imaging, brain and cognitive sciences, earth and environmental science, economics and business, linguistics, physics, political science
- MS specializations: computational methods, statistical methodology, health and biomedical sciences, business and social science

LOCATION

Goergen Institute for Data Science
1209 Wegmans Hall
University of Rochester
Rochester, NY 14627

EDUCATION

GIDS offers three degree programs, a Bachelor of Arts (BA)/Bachelor of Science (BS) program in data science, a Master of Science (MS) in data science, and an advanced certificate in data science, targeted at working professionals. All three programs are flexible and allow students to tailor their degree to their interests. Students are equipped with the technical skills to process and draw conclusions from data, and then trained to apply their conclusions to real-world challenges.

A unique aspect of the undergraduate and MS programs is the capstone course. Through a semester-long capstone project, students get a taste for conducting real-world analytics projects using data provided by

sponsoring organizations. Students work in teams to understand their sponsor's business problem, clean and analyze data, and devise an appropriate solution. GIDS students also have access to nearly 80 affiliated faculty members from across the University of Rochester campus, and can participate in exciting, interdepartmental research collaborations while earning their degrees.

ECONOMIC IMPACT

In 2014, New York State established the Center of Excellence (CoE) in Data Science at the University of Rochester. The CoE in Data Science is dedicated to supporting businesses in New York State through the application of data science methods and tools that solve challenges and deliver critical insights. The Center is funded by the New York State Department of Economic Development's Division of Science, Technology, and Innovation (NYSTAR) and is housed in the Goergen Institute for Data Science. The NYSTAR CoE program helps to drive regional and statewide economic development by supporting basic
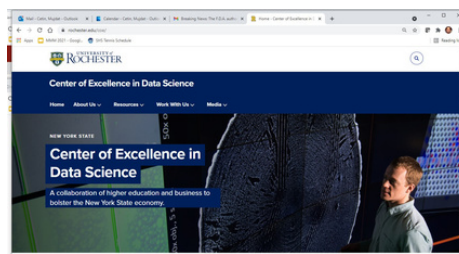
research, training, and technology development in data science. Over the last three years, the Center has generated a direct economic impact of over $110 million and helped create or retain more than 66 jobs. More information can be found on the **Center's web site**.

GOERGEN INSTITUTE
FOR DATA SCIENCE

RESEARCH

GIDS functions as the center for data science research at the University of Rochester (UR) and brings faculty together from across the University together to work on interdisciplinary data science research projects. The Institute is currently involved in multiple collaborative grants in data science, including a grant from the Arnold and Mabel Beckman Foundation on light-sheet microscopy and data science, a National Science Foundation (NSF) PhD training program on augmented and virtual reality, and a joint NSF grant with faculty from Cornell University on foundations of data science.

The Institute also supports six working groups, which explore data science research opportunities around the following themes:
• foundations of machine learning and artificial intelligence
• imaging, optics, and computer/human vision
• life sciences and biomedical data science
• health analytics and digital health
• human-data-system interfaces
• AI augmented learning and work
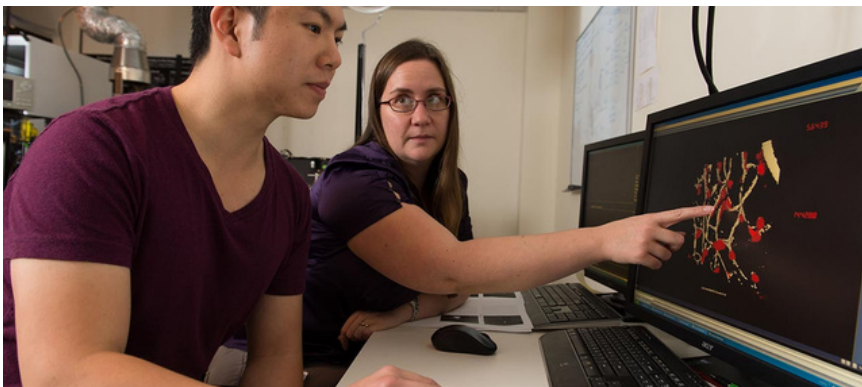
COMMUNITY OUTREACH

Throughout the year, GIDS organizes data science research colloquia for the University of Rochester community. The Institute also boasts a robust alumni network; each semester former data science students return to share their career stories with current students.

GIDS also supports two summer NSF Research Experiences for Undergraduates (REUs). The Computational Methods for Understanding Music, Media, and Minds REU brings undergraduates from across the country to University of Rochester. Under the mentorship of UR faculty members, students explore exciting interdisciplinary research that combines machine learning, music theory, and cognitive science. In the Tripods REU and STEM for All summer research programs, students investigate the mathematical foundations of data science.

GIDS also hosts high school students every summer for a week-long data science immersion. This pre-college program exposes students to data science at a young age, fostering the next generation of data scientists.

Academic
Data Science
Alliance

THE UNIVERSITY OF TEXAS AT EL PASO
DATA SCIENCE
COLLEGE OF SCIENCE
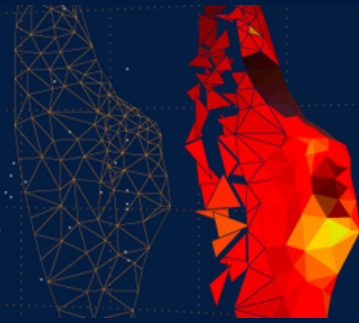
SUMMARY

The Department of Mathematical Sciences at The University of Texas at El Paso (UTEP) houses data science degrees at the bachelors, masters, and doctoral levels. The programs emphasize moving from theory to practice in data science for working in interdisciplinary settings involving data-intensive analysis.

PROGRAM STATISTICS

- Number of faculty: 14
- Number of core staff: 1
- Number of affiliates: 46
- Number of students:
  - BDA Certificate: 54
  - MS: 21
  - PhD: 35

DEGREES, PROGRAMS, SPECIALIZATIONS

- BS in Mathematics with Concentration in Data Science
- MS in Statistics and Data Science
- PhD in Data Science
- Graduate Certificate in Big Data Analytics
- Undergraduate Minor in Data Science

SOCIAL

email: awagler2@utep.edu
web:
www.utep.edu/science/math/academic-programs/graduate/datascience.html

LOCATION

500 W. University Ave.
Bell Hall 311
El Paso, TX 79968

PROGRAM OVERVIEW

This interdisciplinary program is open to qualified individuals from many disciplines who require training on applying data in their fields. The program emphasizes the application and development of


data science methodologies, incorporation of real-world applications through industry and government collaborations, and development of students' professional skills, such as communication and collaboration. The program, designed in consultation with regional industry partners, prepares students to work industries where there are still critical shortages of data science talent.

UTEP DATA SCIENCE COLLABORATIONS

Faculty and students in the data science program regularly collaborate on research projects in partnership with other divisions at UTEP and outside institutions. Some of these are listed below:

- Biostatistics Support for Outcomes-based Medicine (DoD): Annually renewed contract with William Beaumont Medical Center to support research into outcomes-based medicine to address health disparities and optimize care.

- Research Enrichment Core of BUILDing SCHOLARS (NIH): Five year project using data analytics to study factors influencing biomedical research training of underrepresented minorities.
- Border Biomedical Research Center (NIH-BBRC): Long term project that addresses issues of Hispanic Health Disparities unique to the far West Texas Borderplex. UTEP has developed an extensive cadre of collaborators consisting of experts from regional academic, medical, and community partners with the goal to address cancer health disparities that permeate our majority Mexican American population.

Academic Data Science Alliance
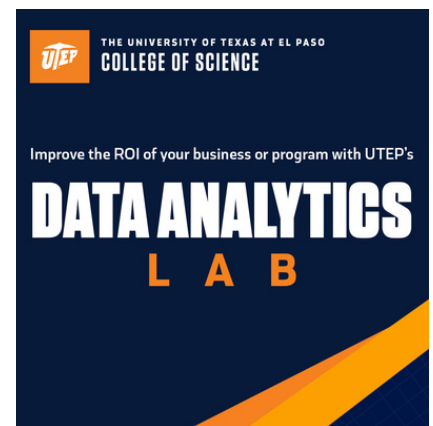
USC Viterbi
School of Engineering
Data Science Program

SUMMARY

Located in the heart of Los Angeles, the University of Southern California is a global center for arts, technology and entrepreneurship that connects students from 64 countries. The USC Viterbi School of Engineering is consistently in the top 10 graduate engineering programs in the U.S. News and World Report rankings.

PROGRAM STATISTICS

- Number of faculty: 21
- Number of technical staff: 10
- Affiliated Faculty: 8
- Students:
  - 750 graduate students
  - 100 undergraduates

SOCIAL

email: datasci@usc.edu

web: datascience.usc.edu

twitter: @USCViterbi

linkedin: www.youtube.com/user/USCViterbi

facebook: www.facebook.com/USCViterbi

instagram: www.instagram.com/USCViterbi

LOCATION

941 Bloom Walk
Los Angeles, CA 90089

PROGRAM OVERVIEW: DATA SCIENCE FOR ALL

Our interdisciplinary data science programs are accessible to students with no programming or computer science background, and are joint programs with other schools at USC. Carefully designed introductory courses on computational thinking and programming enable students to take advanced electives in machine learning, user experience design and strategy, knowledge graphs, and scalable data systems. Students learn to work on teams through experiential learning in course-long projects, hands-on homeworks, class exercises in groups, and capstone classes. Many students take internships during their studies, either in research labs in the engineering school or in local tech companies.

CAREER PATHWAYS

USC is second among all private universities in the number of graduates who join top tech companies. Our graduates pursue careers as leaders of data science teams in diverse sectors including technology, entertainment, health, and policy. Our alumni network spans Fortune 100 companies such as Apple, McKinsey, Capital One, Google, Meta, Amazon, LinkedIn, Bloomberg, and Associated Press, as well as many locally headquartered companies such as Fox Entertainment, Netflix, ESRI, Age of Learning, Warner Bros Entertainment and a myriad others that are part of the vibrant Los Angeles tech landscape known as Silicon Beach.

REMOTE LEARNING

Remote learning is offered for all our graduate data science degree programs through USC's Distance Education Network (DEN@Viterbi). USC is ranked number one in Online Graduate Computer Information Technology Programs (computer science), Online Graduate Engineering Programs, and Online Graduate Computer Information Technology Program (computer science) for Veterans.

FACULTY RESEARCH

Faculty in this program run large data science projects in natural language processing, social network analysis, biomedical data, environmental modeling, and social justice. Some instructors are industry and government practitioners that lead large teams and contribute to community open source software projects.

USC University of Southern California

Academic Data Science Alliance

USCViterbi
School of Engineering
Data Science Program

CORE PROGRAMS

We offer programs at all levels that enable students to learn about a range of topics in data science. Students with previous coursework or experience in data science can work with advisors to design customized curricula.

USC undergraduates can enlist in the Progressive Degree Program (PDP) which enables them to apply undergraduate coursework toward the completion of a USC master's degree in as little as one additional year.

Graduate programs:
- **Master of Science in Applied Data Science**
- **Master of Science in Cyber Security Engineering**
- **Graduate Certificate in Data Science Foundations**
- **Graduate Certificate in Applied Data Science**

Undergraduate programs:
- **Bachelor of Arts in Data Science**
- **Foundations of Data Science Minor**

INTERDISCIPLINARY JOINT DATA SCIENCE MASTERS PROGRAMS

We offer several master's degree programs that enable students to acquire data science skills in the context of a particular discipline. These programs are designed to gently introduce students to programming and computing. Students take courses in computing and data science in the Viterbi School of Engineering as well as courses on quantitative methods and advanced topics in that discipline. All these programs are offered online through our #1 ranked DEN@Viterbi blended platform, enabling students from around the world to complete their graduate degree from a distance while balancing work, life and family.

- **Master of Science in Communication Data Science**, joint with the USC Annenberg School of Communication and Journalism
- **Master of Science in Spatial Data Science**, joint with the USC Dornsife College of Letters, Arts, and Sciences
- **Master of Science in Public Policy Data Science**, joint with the USC Price School of Public Policy
- **Master of Science in Healthcare Data Science**, joint with the USC Keck School of Medicine
- **Master of Science in Environmental Data Science,** joint with the USC Dornsife College of Letters, Arts, and Sciences

THE DATA SCIENCE COMMUNITY AT USC

Our USC Center for Knowledge-Powered Interdisciplinary Data Science (CKIDS) holds semester-long DataFest events that enable students to participate in real-world data science projects advised by faculty across the campus. The GRIDS data science student association provides targeted tutorials and invited presentations from industry on practical applications.

USC CKIDS
Center for Knowledge-Powered Data Science
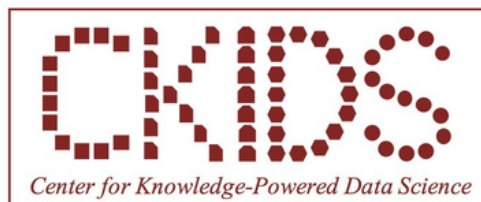
School of Data Science

THE UNIVERSITY OF TEXAS AT SAN ANTONIO

SUMMARY

The School of Data Science (SDS) is a cornerstone of UTSA's vision to reach new levels of excellence in serving our students and society. At SDS, we are inspiring and preparing a generation of diverse data scientists to make the world more equitable, informed, and secure.

PROGRAM STATISTICS

- Number of faculty: 25
- Number of core staff: 5 with 5 additional funded positions to be filled
- In addition to the 26 Core Faculty, 8 faculty members make up the SDS Faculty Council, and all faculty will be invited to affiliate beginning in spring 2022
- Number of students: 351

DEGREES, PROGRAMS, SPECIALIZATIONS

- MS in Artificial Intelligence
- MS in Computer Science – Data Science Concentration
- MS in Data Analytics
- MS in Statistics and Data Science
- PhD Applied Statistics
- UTSA also offers five BS degrees in scope of data science and AI
- Certificates in:
  - Data Science – Graduate
  - Data Science – Undergraduate
  - Gen AI
  - SDS also offers bootcamps in Coding and UX/UI

LOCATION

506 Dolorosa Street
San Antonio, TX 78204

A NEW HOME

In January 2023, the School of Data Science will occupy the new San Pedro I building on UTSA's expanding campus located alongside the San Pedro Creek Culture Park and at San Antonio's urban core. The change to the cityscape will be lasting, but the truly enduring nature of the building will be the incredible work done within its walls when top data researchers innovate with students to make our world more equitable, informed, and secure. In addition to eight data science degree and certificate programs, the School of Data Science will house 12 research centers/institutes/labs.

MATRIX AI CONSORTIUM

The School of Data Science is collaborating with MITRE Corp. to introduce artificial intelligence (AI) principles and practices to UTSA students across all academic disciplines. The goal of *Gen AI* is to help students discover AI functions and to use it as an effective tool.

OPEN CLOUD INSTITUTE

The School of Data Science is home to the Open Cloud Institute to further support its vision to become an industry leader in education and research with a local and global impact.

SOCIAL

email: datascience@utsa.edu
web: provost.utsa.edu/sds
twitter: @utsadata
instagram: @utsadata
facebook: @utsadata
linkedin: www.linkedin.com/company/utsadata

Academic Data Science Alliance

DATA UTAH

School of Computing

UNIVERSITY OF UTAH

SUMMARY

Data science is a rapidly developing discipline that is interplaying with almost every other program or effort at the University of Utah. This interaction is aimed to be collaboratory, and is organized through an interdisciplinary center and a growing variety of academic programs.

PROGRAM STATISTICS

- Number of faculty: 14
- Affiliated Faculty: 13

DEGREES, PROGRAMS, SPECIALIZATIONS

- Bachelors of Science in Data Science
- MS and PhD in Computing - Data Management and Analysis Track
- Professional Masters in Science and Technology - Computational and Data Science

SOCIAL

email: datasci@utah.edu
web: datascience.utah.edu
twitter: @UUDataScience

LOCATION

Salt Lake City, UT, USA

UTAH CENTER FOR DATA SCIENCE

As data science is becoming pervasive, and not limited to any field or subsets of fields, the research and service activities are organized through the Utah Center for Data Science. This center leads, organizes, and manages data science resources and research efforts at the University of Utah. Its members advance the fundamental principles and prctice of data science through research, applications, and community engagement. It hosts a number of regularly occurring events that promote and build engagement in data science.

DATA SCIENCE EDUCATION AT UTAH

While there are many ways data science integrates into academic programs, the central most elements are run out of the School of Computing (https://www.cs.utah.edu/datascience/). These include programs for undergraduates and graduates, involving degrees and certificates. The certificates aim to make these skills accessible to all students with all backgrounds. The focused degrees aim to educate students at the highest skill levels to allow them to specialize and be prepared for any task or career building on their skills in data science.

SCIENTIFIC COMPUTING AND IMAGING INSTITUTE

Data science plays an essential role in many organizations at Utah, a leading one is the Scientific Computing and Imaging (SCI) Institute (https://www.sci.utah.edu). It is a world leader in scientific computing, in biomedical computing, visualization, and image analysis. Each of these domains interplay with data science elements such as machine learning, data management, interactive analysis, and big data.

DATA UTAH

INTERDISCIPLINARY EXCHANGE FOR UTAH SCIENCE

The INterdisciplinary EXchange for Utah Science (NEXUS) is an interdisciplinary research institute at the University of Utah. NEXUS gathers researchers across the University of Utah who desire to collaborate on research projects concerning society's grand challenges, including (1) social disparities & physical / mental health, environment, (2) climate change, the environment and well-being (3) communicating science to the community (4) suicide/violence (5) families and health and (6) opioids.

This institute sparks collaborations through joint research resources, faculty exchange programs, faculty mentoring programs, interdisciplinary summer workshops, interdisciplinary conferences, pilot grant program and graduate and undergraduate training programs. NEXUS also houses the Wasatch Front Research Data Center, which is a secure research environment serving the urban core of Utah and aligned with network of Federal Statistical Research Data Centers.

Academic Data Science Alliance
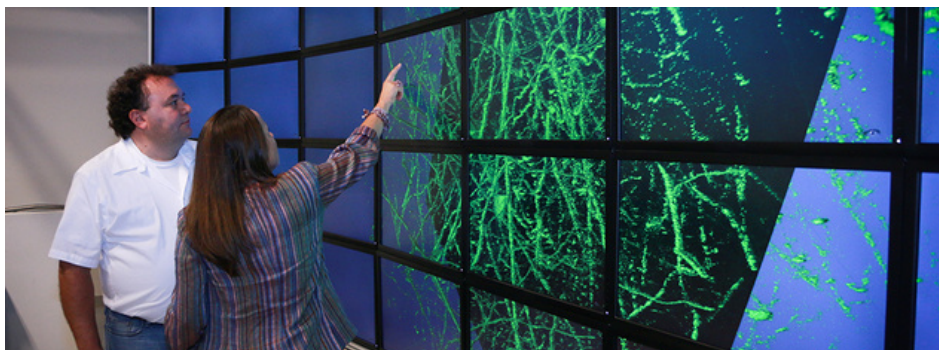
School of Data Science

UNIVERSITY OF VIRGINIA

UNIVERSITY OF VIRGINIA — SCHOOL of DATA SCIENCE

SUMMARY

The University of Virginia School of Data Science—the first of its kind in the nation—is guided by common goals: to further discovery, share knowledge, and make a positive impact on society through collaborative, open, and responsible data science research and education.

DEGREES, PROGRAMS, SPECIALIZATIONS

- M.S. in Data Science (full-time residential and part-time online formats)
- Dual Degrees (MD, PhD, MBA)
- Undergraduate Minor in Data Science
- Professional Programs (Non-Degree)
- Students enrolled in data science related programs (year 2021):
  - 250 MSDS students
  - 200 data science minor
- Certificates in:
  - Deep Learning: A Hands On Approach
  - Data Science for Business Strategy
  - Navigating the Path from Data To Policy

PROGRAM STATISTICS

- Number of Faculty: 23
- Number of Research Staff: 10
- Affiliated Faculty: 17

LOCATION

Charlottesville, VA

CHAMPIONING EQUITY

Our society faces immense challenges related to the COVID-19 pandemic, systemic racism, and broad economic uncertainty. At the School of Data Science, our mission remains clear. From our inception, we have sought to center ethics, inclusion, and transparency in all we do.

CREATING DIVERSE PROGRAM PIPELINES

In addition to partnering with K-12 schools to ensure a diverse candidate pool, we are building relationships with minority-serving institutions to attract diverse students and prospective faculty. We are also increasing our presence at conferences and professional associations focused on serving underrepresented communities.

FOSTERING INCLUSIVE EXCELLENCE

We have established annual merit-based fellowships for an incoming underrepresented graduate student of color or first-generation female. We are also committed to supporting research projects that focus on DEI.

WELCOMING TOP TALENT

In 2021, we hired an Associate Dean of Diversity, Equity, and Inclusion to help to establish a diverse, equitable, and inclusive culture from the ground up. We also launched a Data Activist in Residence position specializing in implicit bias, artificial intelligence ethics, and best-practice criminal justice.

DATA AND OUR COMMUNITY

At the UVA School of Data Science, we are committed to open and responsible data science for the public good while supporting the needs of our local, regional, national, and global communities.

Researchers at the School of Data Science have used data from hospital monitors to better treat patients, credit card data to prevent fraud, and

transportation data to reduce accidents. All of this requires cutting-edge knowledge of data science tools and techniques, as well as the ability to put them into practice. That's why we collaborate with community, industry, and government to connect our work with theirs. Community engagement programs range from Capstone Research projects with community partners focused on the public good to Code for Charlottesville, supporting local nonprofits through the power of data science, computing and student creativity.

UNIVERSITY OF VIRGINIA

SCHOOL *of* DATA SCIENCE

RESPONSIBLE DATA SCIENCE

Our field may be nascent, but its potential for producing nefarious outcomes has already been well documented. We are committed to teaching and practicing responsible data science, with the common good in mind. As the first school of data science in the country, we have an obligation to be a leader in the field and in education. That leadership is built upon ethical partnership with all stakeholders, from schools and departments at UVA, to other universities world-wide, to the private sector, to government, to nonprofits, and to the general public. We are also leading our institution in furthering open scholarship and are establishing an institution-wide Open Scholarship Working Group.

FURTHERING DISCOVERY

The School of Data Science pursues high-impact research to further discovery, share knowledge and transform society. Through their research, our faculty and students are building a better world in a variety of ways.

Democracy: Investigating how terrorist groups recruit women through propaganda and examining risk for extremist violence.

Education: Helping economically disadvantaged, underrepresented populations pursue pathways that have a higher probability of leading them to success.

Business: Discovering what makes a job interview successful for both the candidate and the recruiter and learning how to mitigate bias in the recruiting process.

Health & Medicine: Securing high-performance computing equipment and personnel to allow collaboration across the university on brain science research concerning autism, participating in translational biomedical data sciences, and more.

Cybersecurity: Detecting broad spectrum cyber threats almost immediately after they are launched—research made possible through a grant from the Department of Defense.

Environment: Using NASA data collected aboard the International Space Station to examine and develop responses to climate change in the Shenandoah National Forest and beyond.

SOCIAL

email: datascience@virginia.edu
web: datascience.virginia.edu
twitter: @uvadatascience
linkedin: wlinkedin.com/school/uvadatascience
facebook: facebook.com/uvadatascience
instagram: instagram.com/uvadatascience

Academic Data Science Alliance

eScience Institute

SUMMARY

The eScience Institute was founded in 2008 with the mission to empower researchers and students in all fields to answer fundamental questions through the use of large, complex, and noisy data. From inception, the eScience Institute has functioned beyond departmental and college walls to exemplify the interdisciplinary breadth and complex dimensions of data science. Today, the Institute has grown to a mature organization with sustained positive impact through our education, research, and community building programs.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Masters in Data Science
- Undergraduate Minor
- 17 Graduate and 8 Undergraduate Specializations

PROGRAM STATISTICS

- Number of faculty: 14
- Number of staff: 24
- Affiliated Faculty: 130

LOCATION

eScience Institute
WRF Data Science Studio
Physics/Astronomy Tower (PAT), 6th Floor
3910 15th Ave NE
Seattle, WA 98195-1570

PROGRAM OVERVIEW

The University of Washington eScience Institute, one of the nation's first university data science institutes, grew out of the Moore-Sloan Data Science Environment effort which focused on identifying and tackling impediments to the broad and sustainable adoption of data-intensive discovery. From inception, the eScience Institute has functioned beyond departmental and college walls to exemplify the interdisciplinary breadth and complex dimensions of data science. Today, the Institute has grown to a mature organization with sustained positive impact on the UW community through our education, research, and community building programs. Key eScience personnel now hold leadership roles in the construction, commissioning and operations of federally-funded research initiatives such as Pangeo, the Rubin Observatory, and Interactive Oceans as well as NSF-funded programs including CloudBank and the West Big Data Innovation Hub.

INSTITUTIONAL SCOPE, STRUCTURE, & ACTIVITIES

The institute comprises a community of innovators in the development and application of the techniques, technologies, and best practices of data science. The eScience conception of data science has always considered the full data science workflow, from project conception to data-driven discovery and every step in between. With expertise in advanced statistical and computational techniques including artificial intelligence, machine learning, database management, visualization, and research software engineering, our staff Data and Research Scientists have energized the WRF Data Science Studio into a nexus of data science activity on campus and virtually. We serve campus data science needs and propel data science research forward via multiple channels from open office hours to quarter-long engagements like the Incubator and Data Science for Social Good programs to multi-year collaborative grants and centers. Our faculty, postdoctoral fellows, and graduate students are engaging in cutting edge research pushing the methodology envelope and making new discoveries across a diversity of fields.

DATA SCIENCE INCUBATORS

The Data Science Incubator enables new science by bringing together data scientists and domain scientists to work on focused, collaborative projects. Each fall, we invite proposals for a quarter-long data-intensive research collaboration focusing on extracting insight from large, noisy, and/or heterogeneous datasets. Over the past 8 years, we have supported 50 projects with collaborators from 29 different departments. Previous incubator projects have led to long-term collaborations, publications, and grant awards.

Launched in 2015, the UW's Data Science for Social Good (DSSG) summer program partners Student Fellows with Data Scientists from the eScience Institute and Project Leads from academia, government, and the private sector to find data-intensive solutions to pressing societal challenges. Previous projects have used diverse methods to address topics such as public health, homelessness, disaster response, voting

rights and transportation. Keystones of the DSSG program include project-based discussions and training around data science ethics, human-centered design and stakeholder collaboration.

HACKWEEKS

The hackweek model has emerged within the data science community as a powerful tool for fostering the exchange of ideas in research and computation. In contrast to conventional conferences or workshops, hackweeks are intensive and interactive, facilitated by 3 core components: tutorials on state-of-the-art methodology, peer-learning, and on-site project work in a collaborative environment. This setup is particularly powerful for sciences that require not only domain-specific knowledge, but also effective computational tools and methodologies. The eScience Institute has extensive experience (check out this toolkit) developing and facilitating hackweeks focused in particular domains, e.g. Neurohackademy and the Electrochemical Society HackWeek, as well as around particular datasets, e.g. the ICESat-2 and SnowEx Hackweeks.

Academic Data Science Alliance

data science @ uw

UNIVERSITY OF WISCONSIN - MADISON

SUMMARY

Data is reshaping our world, and data science @ uw strives to bring the power of fundamental and applied data science to all fields of study at UW-Madison and beyond. We are committed to fostering an inclusive culture in data science that fuels creativity and discovery.

PROGRAM STATISTICS

Over 500 undergraduate Data Science majors. Over 100 undergraduate Data Science certificate students. Hundreds of other students are enrolled in the rich array of data science graduate education offered at UW-Madison.

DEGREES, PROGRAMS, SPECIALIZATIONS

- Undergraduate Major in Data Science
- MS Statistics: Data Science
- MS Industrial Engineering: Systems Engineering and Analytics
- Engineering Data Analytics Online Program
- MS Business Analytics
- Master of Science in Psychology: Data Science in Human Behavior
- Master of Science in Educational Psychology: Learning Analytics
- Undergraduate Minor (certificate) in Data Science

LOCATION

American Family Insurance Data Science Institute, 447 Lorch St., Madison, WI 53706
Data Science Hub, 330 N Orchard St., Madison, WI 53715
School of Computer, Data & Information Sciences, 1210 W. Dayton, Madison, WI 53706

AMERICAN FAMILY INSURANCE DATA SCIENCE INSTITUTE

Established in 2019, the American Family Insurance Data Science Institute (DSI) performs cutting-edge research, bridging theory and practice, in collaboration with people across the UW-Madison campus and around the world. Three primary functions drive the work of DSI: Innovation, Translation, and Collaboration. Cutting-edge research projects bring people together to develop innovative data science methodology. Teams of data scientists and domain research scientists translate frontier research methods into practical, adaptable resources. Facilitators, working through the Data Science Hub, connect researchers from all disciplines with available data science resources, experts, and efforts. DSI fosters an inclusive and responsible culture in data science that fuels creativity and discovery.

DSI is central to UW-Madison's strategic priority to grow its research enterprise and expand its global impact, supporting the scholarship of faculty, staff, and students. Our strategy is informed by the Wisconsin Idea: Research and education should influence people's lives beyond the boundaries of the classroom, lab, and campus.

DATA SCIENCE HUB

The Data Science Hub in the Wisconsin Institute for Discovery collaborates closely with the DSI to provide training and implement data science practices into research across campus. The Data Science Hub executes its mission of community engagement and learning opportunities for researchers across campus through a variety of services. The Data Science Hub hosts trainings and workshops around fundamental data science and computational skills, to help researchers learn to reproducibly write software and analyze data. It offers consultations with data science facilitators who can recommend
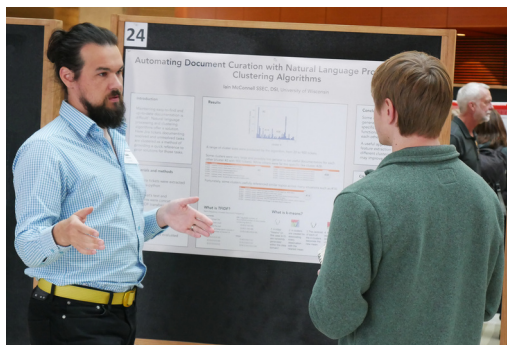
learning pathways and project strategies, and liaise contacts with collaborators and data science experts. The Data Science Hub organizes seminars and events, including UW-Madison's annual Data Science Research Bazaar, where researchers and data scientists from different disciplines and industries have opportunities to share their work, collaborate, and discuss their data science interests.

SCHOOL OF COMPUTER, DATA & INFORMATION SCIENCES

The School of Computer, Data & Information Sciences (CDIS) was launched in 2019 and brings together the top-ranked departments of Computer Sciences, Statistics, and the Information School. CDIS is lighting the way forward for research and discovery, magnifying the power of medicine, engineering, agriculture, business, and more. CDIS is educating versatile, multi-talented graduates across all majors on campus—who form a uniquely prepared talent pipeline that is driving economic growth in the region and beyond.

CDIS collaborates across campus, regionally, and nationally to produce cutting-edge, transformative research, educate leaders and critical thinkers, and accelerate innovation that tackles societal issues. Our departments are renowned for the groundbreaking research that our faculty, staff, and students do. From undergraduate and graduate degrees to certificates, professional development, and workshop opportunities, CDIS helps students earn the tech skills they need to be successful leaders and strategic thinkers in their fields.

OTHER DATA SCIENCE-FOCUSED ORGANIZATIONS AT UW-MADISON

Department of Biostatistics and Medical Informatics (BMI)
BMI advances data science for medicine and public health through collaboration with biomedical scientists.

Institute for Foundations of Data Science (IFDS)
IFDS, funded by NSF's TRIPODS program, focuses on fundamental, theoretical issues in data science. IFDS faculty are involved in applications-focused projects.

Machine Learning and Optimization Research Consortium (MOR)
MOR works with industry partners to solve machine learning and optimization challenges in data analytics, healthcare systems, manufacturing, transportation, and other commercial applications.

Machine Learning for Medical Imaging (ML4MI)
ML4MI fosters interdisciplinary collaboration between machine learning experts and medical imaging researchers in order to solve problems in medical imaging.

MADLab
MADLab is a University Center of Excellence supported by the Air Force Office of Scientific Research and the Air Force Research Lab to develop the next generation of machine learning theory, algorithms, and applications.

Wisconsin Institute for Discovery (WID)
WID's expertise in data science spans disciplines, with the goal of developing end-to-end strategies for data collection, analysis, management, privacy, security, and decision-making.

Data Science Instiute

VANDERBILT UNIVERSITY

SUMMARY

The Vanderbilt Data Science Institute accelerates data-driven research, promotes collaboration, and trains future leaders. The institute brings together experts in data science with leaders in all academic fields to spark new discoveries. The institute educates students in data science to become leaders in industry, government, academia and the nonprofit sector.

PROGRAM STATISTICS

- Number of teaching faculty: 14
- Number of professional staff: 6
- Number of affiliated faculty: 224
- Number of students: 82

DEGREES, PROGRAMS, SPECIALIZATIONS

- Master of Science in Data Science
- Undergraduate Minor in Data Science

LOCATION

Vanderbilt University
17th & Horton Building
1400 18th Ave S, Nashville, TN 37212

DATA SCIENCE EDUCATION

M.S. in Data Science - Our two-year data science degree boasts a customized data science-specific curriculum taught by an interdisciplinary faculty from Medicine, Engineering, Arts and Sciences, Education, and Business. Our program is 46% female and we are committed to enabling gender equality in data science. The Owen Graduate School of Management's highly-ranked Career Management Center supports our students' career goals. Our program is rigorous, practical, experiential, team-oriented, and student-focused. Our Data Science Team attracts a continual flow of real-world projects for students to participate in both in and out of the classroom. See our promo video.

Data Science Minor - Vanderbilt's trans-institutional data science (DS) minor provides students of all disciplines an accessible and well-rounded foundation in DS, providing context and experience with the tools of our increasingly data-driven society.

This 19-credit hour interdisciplinary program covers foundational DS skills, from programming to statistics, interweaving ethical considerations of data use and interpretation. With experiential learning opportunities, the minor provides a solid foundation for future professions or graduate study in any field that uses data.

INTERDISCIPLINARY WORK

At the DSI, we engage in projects across the academic and industry spectrum. From archaeology to medicine to literature, students have the opportunity to participate in new and ongoing collaborations, or to explore their own. Every Friday the DSI hosts AI Fridays, where researchers drop by for a Deep Dive, hour-long in-depth conversation with data scientists and faculty on a topic of interest. Recent deep dives have included using transformers to study special education classroom interactions to the application of AI in design and engineering problems in industry. Students have multiple opportunities to explore applications of data science. Our Masters students complete a self-directed Capstone project which is often in partnership with faculty in various disciplines or industry partners. In our Data Science for Social Good Programs, students have an opportunity to make a substantial contribution by applying data science skills, while gaining the opportunity to solve problems across a broad range of areas. Undergraduates can participate in the DSI Summer Research Program to pursue data science-related research with a faculty sponsor.
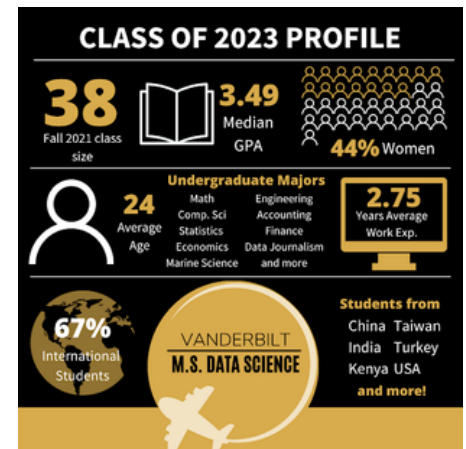
PROFESSIONAL DATA SCIENCE EXPERIENCE

"I have learned that communication is super important for a future data scientist. In the program we practice a lot on how to communicate, and now I understand that your team members all have the same goal - completing the project in the best way possible. It opens you up to different opinions and ideas."

Marina Yue

Bachelor of Arts
Statistics
University of Florida

VANDERBILT UNIVERSITY® | Data Science Institute
Discovery through data.

The Data Science Institute at Vanderbilt is actively engaged in data science projects and students are an integral part of those teams. As one of its core groups, the DSI has a professional data science team that engages with researchers, industry, and not-for-profits in a dozen or more projects annually. Each is run using agile processes and best practices. By building teams of staff data scientists, Masters students and undergraduates, the DSI is able to both support the development of cutting-edge tools and approaches, provide real-world experience, and give students experience in fully participating as a team member in a high-functioning data science team.

This emphasis on real-world experience and best practices of team data science is woven in throughout the curriculum. The Data Science in Practice class is now in its third year of teaching students that "Data Science is a Team Sport'' with industry best practices using real-world corporate projects. Project classes allow for a focused semester-long team engagement while also introducing new tools.

INDUSTRY PARTNERSHIPS

The DSI is actively engaged with industry through consulting, collaborative data science projects, course-based projects, and capstones. Through its Data Science Team, the DSI provides consults on data science and AI solutions to industry ranging from new startups to multinational corporations. Partners include AllianceBernstein, AWS, Bridgestone, Asurion, The General Insurance, and others.

Students have an opportunity to participate in projects with direct impact on businesses by joining data science teams that tackle scoped projects with our partners. They apply what they've learned in classes while developing new skills that can only be developed working on a team in a business environment. By graduation, students have had the opportunity to work with multiple companies in a real-world team environment, and are ready to step into data roles and immediately contribute.
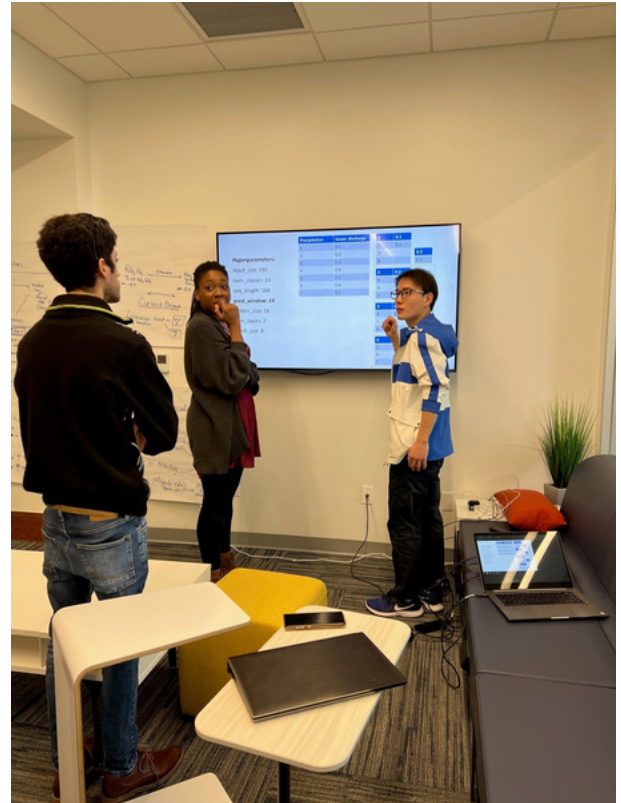
Academic Data Science Alliance

Center for Applied Data Science

WINSTON-SALEM STATE UNIVERSITY

SUMMARY

Founded in 1892, Winston-Salem State University is a public Historically Black College and University (HBCU) and an University of North Carolina (UNC) institution, where 76% of the student population are African Americans or Hispanic/Latino, 73% are female, and 23% are first generation college students.

PROGRAM STATISTICS

- Number of Faculty: 1
- Number of technical staff: 3
- Affiliated Faculty: 6

SOCIAL

email: debd@wssu.edu
web: www.wssu.edu/cads
linkedin: www.linkedin.com/company/wssu-cads
youtube: www.youtube.com/channel/UCkRpCwBL-Ml6lQw13-3dJqA

LOCATION

Winston-Salem, NC, USA

PROGRAM OVERVIEW

The Center for Applied Data Science (CADS) at WSSU is an institution-wide initiative with the goal of fostering research and education in data-driven knowledge discovery. CADS aims to bring together computer scientists and domain scientists with complex Data Science problems to promote and accelerate data-intensive discovery and education. Ongoing research projects are in the areas of pharmacoengineering, mobile crowdsensing, spatial justice and social mobility, healthcare management, and music's biophysical influence on the human body. The center also focuses on leveraging the power of data to help change lives and impact our local community, while also fostering greater diversity and inclusion in the rapidly growing field of data science. The center received its inaugural funding of 1.5M in the Fall of 2020 for three years funded by the UNC Research Opportunities Initiative (ROI). Currently, 30 faculty, students, and staffs are supported by CADS under the leadership of Debzani Deb, the founding director of CADS.

FACULTY ADOPTER AWARDS

To better integrate data science into WSSU's full curriculum, the center started Faculty Adopter Awards which are designed to support enthusiastic WSSU faculty who are willing to infuse data science into their courses across various disciplines and are capable of quantitatively and qualitatively assess the impact of their interventions. The FY 21 and FY22 cohort of 10 awardees developed 10 relevant course modules in data science and integrated them into existing courses offered at various disciplines such as Arts, Education, Mathematics, Sports Studies, Psychological Sciences, Justice Studies, and Healthcare Management. As part of this effort, 150 students received exposure to data analytics and gained hands-on experience on using data science tools and techniques.

CADS RESEARCH PROJECTS

CADS focuses on exploring interdisciplinary research projects and bridging the gap between domain experts and computer scientists to maximize the power and practicality of data science. Several current research projects focus on eliminating healthcare disparities and advancing health equity by creating an early recommendation system and investigating the prominent factors that cause such disparity. Another project investigates the application of mobile crowdsensing data to improve user experience, facility management, and resource usage. Data science is also being explored to make therapeutic predictions by mining publicly available gene expression data to repurpose current FDA-approved drugs. Another CADS researcher is applying data science in identifying the human body's response to various genres of music to aid individuals' decisions to listen to a particular genre to obtain the desired health outcome. CADS Researchers are also exploring geographic, demographic, and socioeconomic variables to help communities address spatial injustices.

Center for Applied Data Science

WINSTON-SALEM STATE UNIVERSITY

DEGREES, PROGRAMS, SPECIALIZATIONS

CADS offers an interdisciplinary minor in data science, facilitates the gradual infusion of data science into the university-wide curriculum across diverse disciplines through Faculty Adopter Awards, supports existing graduate certificate in data analytics, and aims to offer a MS in Data Science in near future.

Through a joint effort of CADS and the Department of Computer Science at WSSU, a Data Science Minor is available to all WSSU students, beginning in Fall 2021. The Data Science Minor requires 6 courses (18 credit hours) – including 3 foundational courses in data science and statistics and 3 approved elective courses offered at various academic departments. Through participating in the minor, students will gain skills in data acquisition and management; data analysis; solving real world data-based problems, interpretation and communication of data; and understanding the social and ethical implications of work in this field.

DEI AND OUTREACH

We expect that our various educational efforts and the research student mentorship efforts will expose women, minority, and underrepresented students to data science content and skills that may influence their choice about further education and future careers and therefore increases the size and diversity of the future STEM workforce while changing lives and communities for better.

In fact, we plan to begin raising awareness of data science careers long before students even enroll at the university. The Center will contribute to K-12 outreach activities, including camps at local middle and high schools, to engage students and teachers in the possibilities of data science and how the field can be integrated into K-12 learning.

COMMUNITY COLLABORATION

While CADS' work is occurring in academic settings, the practical outcomes will have a lasting and far reaching impact on the local community. We are eager to team with local businesses that might benefit from data science as they pursue investment and commercialization of their products. We also invite academia, industry, nonprofit, and government agency participation in our annual symposium and monthly seminars to increase awareness of data science. The Center gives us all a chance to do something very special together – prepare our young people for a bright future, improve lives through work that also bolsters the economy, and advance the cause of racial justice and equity throughout the Triad and North Carolina.

Yale

DEPARTMENT OF STATISTICS AND DATA SCIENCE

YALE UNIVERSITY

SUMMARY

The Yale Department of Statistics and Data Science is committed to furthering education and research in the rapidly growing field of data science, ranging from its foundations in mathematics, statistics, and computing to its applications to cross disciplinary research addressing fundamental problems in every area of inquiry and endeavor.

PROGRAM STATISTICS

- Number of faculty: 14
- Number of core staff: 4
- Number of students: 15

DEGREES, PROGRAMS, SPECIALIZATIONS

- BA in Statistics and Data Science
- BS in Statistics and Data Science
- MA in Statistics
- MS in Statistics and Data Science
- PhD in Statistics and Data Science
- Undergraduate Certificate in Data Science

SOCIAL

email: sds_chair@yale.edu
**web**: sds.yale.edu

LOCATION

24 Hillhouse Ave.
New Haven, CT 06511

PROGRAM OVERVIEW

The Statistics and Data Science (S&DS) department offers Ph.D. and masters programs, an undergraduate major, and a new certificate program open to undergraduates from any major. These educational programs are high priorities, and the department has been recognized for teaching excellence. We strive to maintain close personal attention to students as much as possible while teaching some of the largest and most popular courses on campus. Research topics pursued within the department span a range from foundational topics in probability and statistical theory to the most current challenges in machine learning and optimization. Such research directions include high-dimensional statistics and machine learning, theoretical computer science and algorithms, optimization, information theory and its connections with probability and statistics, covariance matrix estimation, network analysis, function estimation and neural networks, numerical analysis, signal and image processing, empirical geometry of data, estimation of causal relationships, statistical computing and the practice of data analysis, and the responsible and ethical use of data, algorithms, and machine learning.

Yale *Department of Statistics & Data Science*

Anscombe's Quartet (1973)
by Frank Anscombe, Founding Chair of the Department

The *first* Statistics and Data Science NFT

Ref: NFT/Yale/S&DS/1 (pilot)

OUR VIEW OF DATA SCIENCE

The department takes a broad and multi-faceted view of data science, encompassing the entire life cycle of data, from its specification, gathering and cleaning, through its management and analysis, to its use in making decisions and setting policy. Areas of application under active development come from a wide variety of fields and problems, such as genetics and computational biology, structural biology and cryo-EM microscopy, new computational frameworks for neuroscience, autism research, evolutionary relationships among species, large-scale multi-agent decision-making problems, the design of field experiments in political science, and algorithmic interventions to address systemic forms of discrimination. This work is done in collaboration with many organizations that include other academic departments at Yale such as Mathematics, Political Science, and Computer Science, groups in the Yale School of Medicine such as the Center for Outcomes Research and Evaluation and the Child Study Center, the Institute for Network Science, Quantitative Biology Institute, Institution for Social and Policy Studies, and the Cowles Foundation for Research in Economics.
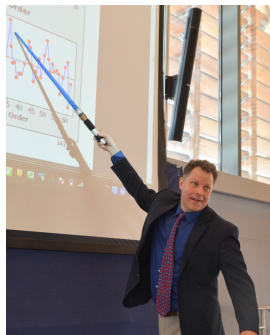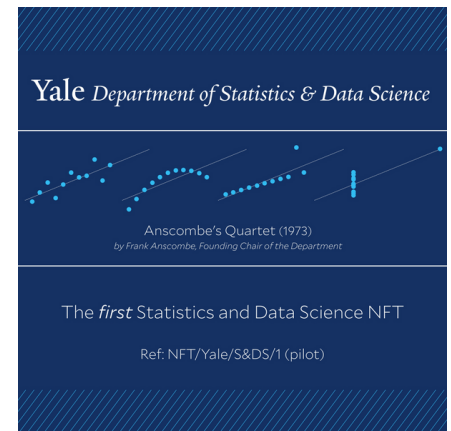
DATA SCIENCE AT YALE IS GROWING

Data science at Yale is in a period of tremendous growth. The S&DS Department has hired eight faculty in the last four years, several with joint appointments with other departments. In the undergraduate major, for example, around 10 years ago, the number of graduating seniors each year would average around 3, while in recent years our program has grown so that last year the number of graduating seniors majoring in S&DS was over 60. In 2018, Yale announced that Data Science would be one of its top five priority areas for new investments in science. The initiatives proposed then by the University Science Strategy Committee included establishing a university-wide Institute for Data Science. Following up on this initiative, the "Kline Tower Institute for the Foundations of Data Science" is now forming and will foster collaboration between many different components of the university, bringing together faculty and students across the sciences, social sciences, arts and humanities, and professional schools such as the medical school, law school, and school of the environment.

A NEW HOME FOR S&DS

Since 1963, the department has enjoyed working in a lovely national historic landmark, the former home of geology professor James Dwight Dana. Perhaps encouraged by this setting, in which many of the faculty offices were once bedrooms and department members would gather in what was once a living room, the department has always been very much like a family, with a warm and friendly environment that we have always valued. With our recent expansion, we have overflowed this home (as shown in the photo below) to expand into parts of two other buildings just down the street and just across the street. Currently, planning is underway for the department to be reunited in a single building, the Kline Tower (photo above) in the summer of 2023. The S&DS Department will occupy 3 floors of Kline Tower, and the new Kline Tower Institute will be on the floor immediately above S&DS.

Academic Data Science Alliance

We would like to thank the following, additional, founding institution:

- Carnegie Mellon University

In addition to members listed in the previous sections, we would also like to thank our individual ADSA members, representing the following institutions:

- Amherst College
- California State University, Fresno
- Embedded Intelligence
- George Washington University
- James Madison University
- Rice University
- Tel Aviv University
- University of California, Irvine
- University of Pittsburgh

Academic Data Science Alliance

www.academicdatascience.org