VANDERBILT V UNIVERSITY ®

Data Science Institute

DATA SCIENCE SYMPOSIUM

Friday, April 26, 2019
Vanderbilt University – Student Life Center Ballroom

DSI INFORMATION

- Promote **data-intensive research** and **quantitative collaboration** across Vanderbilt.

- Provide a forum to study the **societal impact** of 'big data'.

- Facilitate **educational programs** for the community at large.

- Cultivate new and existing **partnerships** with government, industry, and non-profit entities.

Data Science Master's Degree

Training the next generation of data-driven research specialists.

An interdisciplinary program with a balanced curriculum in statistics, computational algorithms, and communication.

Pillars of the Program

computational pipelines for high-dimensional complex data

machine learning algorithms for prediction & decision making

statistical models and data visualization

ethics, policy, & privacy

teamwork & professional skills

www.vanderbilt.edu/datascience

datascience-grad@vanderbilt.edu

@VUDataScience

COURSES OFFERED IN:
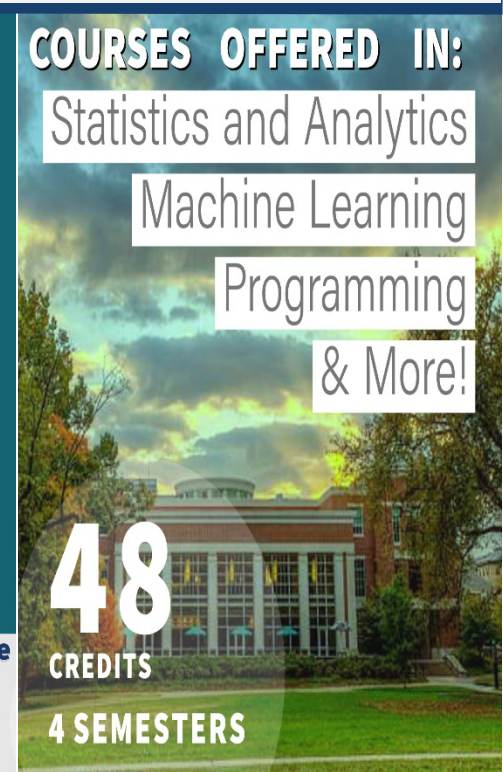Statistics and Analytics
Machine Learning
Programming
& More!

48 CREDITS
4 SEMESTERS

PROGRAM

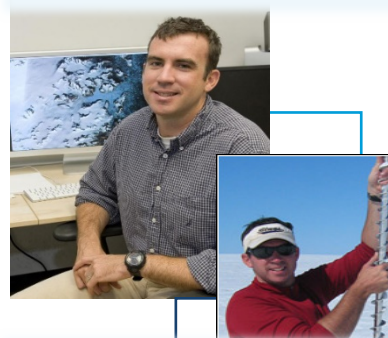|	
8:30 AM	**Graduate Student Poster Session**
9:45 AM	**Welcome & Opening Remarks** Andreas Berlind
10:00 AM	**Short Talks** Hiba Baroud
11:00 AM	**Lightning Talks** Lindsey Fox
12:00 PM	**Lunch**
12:15 PM	**Poster Award Session**
12:20 PM	**Changing Practice of Data Science:** **A Panel Discussion** Jesse Spencer-Smith
1:00 PM	**Terra Cognita: The Silicon Age of Earth Exploration** Ian Howat
2:00 PM	**Coffee Break**
2:30 PM	**Short Talks** Jeffrey Blume
3:20 PM	**Live Deep Learning Demo** Jesse Spencer-Smith

KEY NOTE

DR. IAN HOWAT
PROFESSOR, SCHOOL OF EARTH SCIENCES
THE OHIO STATE UNIVERSITY

Terra Cognita: The Silicon Age of Earth Exploration

The serendipitous convergence of three trends over the past few years has enabled repeat, high resolution terrain mapping at the continental scale; these are the opening of access to large quantities of sub-meter stereoscopic satellite imagery, the creation of efficient, open source image processing software, and the availability of petascale high performance computing (HPC). Within a matter of months, vast areas at high latitudes and other remote regions have gone from having the poorest topographic data, to among the best on Earth; and these data are freely open to everyone. Here I review progress to date in the pursuit of global, time-dependent terrain mapping and the new science it is enabling. I include some lessons learned from the first forays into continental scale mapping with these data: the ArcticDEM and Reference Elevation Model of Antarctica (REMA) projects. I will highlight some of the novel initial approaches taken to ingest and analyze the large body of information these data provide. Based on this early progress, I will conclude with a few recommendations for enhanced utilization and continued improvement of this transformative new resource for understanding the Earth and how it is changing.

Ian Howat has spent over 20 years studying how glaciers and ice sheets respond to climate forcing on time scales ranging from minutes to millennia and ranging in scale from the formation of individual icebergs to the mass balance of the Greenland and Antarctic Ice Sheets. An expert in airborne and satellite remote sensing, Ian has pioneered efforts to map rapid changes in the polar regions. He has been a participant and leader on numerous science expeditions to Antarctica, Iceland and Greenland, and is the principle investigator of the Reference Elevation Model of Antarctica (REMA) project, producing the first, high-resolution terrain map of a continent. He received the Presidential Early Career Award in Science and Engineering (PECASE) in 2011 and He and his collaborators were awarded the Hyperion Innovation award for their work on the ArcticDEM project in 2017.

Ian obtained a BA in Geology from Hamilton College and a PhD. in Earth Science from the University of California, Santa Cruz. He is currently a Professor in the School of Earth Sciences at Ohio State University and is the Director of Ohio State's Byrd Polar and Climate Research Center.

SHORT TALKS
MORNING SESSION

Challenges and Promise in Precision Medicine
Andrea Ramirez
Assistant Professor of Medicine

Data-Driven Product Development at Google
Jason Schwarz
Quantitative Researcher, Google

From Single-Cell Genomics to Tissue Atlases
Ken Lau
Assistant Professor of Cell and Developmental Biology

Data and Mind: How People Perceive and Process Visual Information
Maithilee Kunda
Assistant Professor of Computer Science

Machine Learning Methods for Energy Reduction in Large Buildings
Gautam Biswas
Cornelius Vanderbilt Professor of Engineering

SHORT TALKS
AFTERNOON SESSION

Decomposing Medicare Spending Growth
Melinda Buntin
Professor and Mike Curb Chair of Health Policy

Gentrification and Access to Transportation in Nashville: A Data-Driven Approach
Jonathan Gilligan
Associate Professor of Earth and Environmental Sciences

Election Forensics Using Voter File Data: The Case of North Carolina
Josh Clinton
Professor of Political Science

Visualization of Social Media Data Using Splunk
Nadine Wondem
Principal Cybersecurity Engineer, Defense Point Security, LLC

CHANGING PRACTICE OF DATA SCIENCE PANEL

Middle Tennessee data scientists discuss advancements that have changed their work.
- **Tim Blass, PhD:** Lead Data Scientist, Financial Services, Digital Reasoning
- **Sharon Chou, PhD:** Principal Data Scientist, Amira Learning
- **Mathilde Granke, PhD:** Senior Research Data Scientist, axialHealthcare
- **Lucas Lukasiak:** Director Healthcare Analytics, Change Healthcare

DEEP LEARNING LIVE

Running code live during a presentation is ill-advised--training a deep learning model is worse still. At the beginning of the symposium we'll begin training deep neural network, with plans to have it fully trained and available by the last presentation of the day. We'll demonstrate how well the model learned, and show results that should be of interest to all attendees.

GRADUATE STUDENT POSTERS

Abin Abraham: Boosted Decision Trees Accurately Predict Preterm Birth with Structure and Unstructured Data from Electronic Health Records

Ky'Era Actkins: Association of inactivating calcium sensing receptor exon 7 SNPs with hypercalcemia-related disease phenotypes

Moyo Ajayi: Real-Time Analysis Gas Emissions during Geyser Eruptions

Gillian Beltz-Mohrmann: Can we ignore baryons in halo modeling?

Mary Lauren Benton: Genome-wide enhancer maps differ significantly in genomic distribution, evolution, and function

Kelsea Best: Environmental migration and land inundation: an agent-based modeling approach

Qinyuan Gu: Examining the prevalence of depression and anxiety symptoms in the Chinese student population at Vanderbilt

Coleman Harris: Statistical Efficacy of Distance Matrix Anonymization

Benjamin Kesler: Dissecting long non-coding RNA regulation with single-cell image analysis

Ray Matsumoto: Predicting Phase-Separation in Molecular Dynamics Simulations of Ionic Liquid-Organic Solvent Mixtures

Megha Patel: Using Data Driven Methods to Estimate Tennessee's Tax Burdens

Daniel Perrucci: Are Mitigation Efforts Effective at Reducing Natural Disaster Risk?

Stephen Robinson: Predicting School-level Performance: A comparison of spatial and non-spatial methods

Chris Tasich: Predicting arsenic contamination in groundwater wells in the Bengal Basin

Hannah Weeks: A novel natural language processing algorithm for medication dose extraction from electronic health records

Patrick Wu: Dynamics of Physician-Scientist Howard Hughes Medical Institute (HHMI) Investigators, 2000-2018

Jin-Zhu Yu: Quantifying Community Resilience Using Hierarchical Bayesian Methods: A Case Study on Recovery from Power Outages

Linda Zhang: Automated Sleep Stage Scoring Using Deep Learning

DATA SCIENCE INSTITUTE

LEADERSHIP

ANDREAS BERLIND
Co-Director

DOUG SCHMIDT
Co-Director

JEFFREY BLUME
Director of Graduate Education

JESSE SPENCER-SMITH
Chief Data Scientist

AMANDA HARDING
Assistant Director

EXECUTIVE COMMITTEE

HIBA BAROUD, Civil & Environmental Engineering
JOSH CLINTON, Political Science
BRAD MALIN, Biomedical Informatics
GREGOR NEUERT, Molecular Physiology and Biophysics
SONYA STERBA, Psychology & Human Development

AFFILIATED FACULTY

Thank you to our affiliated faculty for your continued interest and support. If you would like to be affiliated with the Data Science Institute, please visit: https://www.vanderbilt.edu/datascience/people/affiliates/