GRADUATE STUDIES IN COMPUTATIONAL & DATA SCIENCES

RACKHAM GRADUATE CERTIFICATE IN DATA SCIENCE
(MIDAS)

GRADUATE CERTIFICATE IN COMPUTATIONAL DISCOVERY AND ENGINEERING

PHD PROGRAM IN SCIENTIFIC COMPUTING (MICDE)

September 24 & 25, 2018
Advanced Research Computing

Leading advances in data-intensive and computational research

**COMPUTATIONAL SCIENCE**
The Michigan Institute for Computational Discovery and Engineering (MICDE) focuses on the interdisciplinary development of mathematical algorithms and models on high performance computers.

**DATA SCIENCE**
The Michigan Institute for Data Science (MIDAS) is the focal point for the new multidisciplinary area of data science at the University of Michigan.

**TECHNOLOGY SERVICES**
Advanced Research Computing – Technology Services (ARC-Ts) provides access to and support for the use of advanced computing resources.

**CONSULTING SERVICES**
Consulting for Statistics, Computing and Analytics Research (CSCAR) provides support and training relating to the management, collection, and analysis of data.
Computational Science vs Data Science

- **Computational Science**
  - Laws of Physics
  - Mathematical Models
  - Numerical Algorithms
  - Solutions
  - Simulation Data

- **Data Science**
  - Big Data
  - Algorithms
  - Patterns, Statistics, Smart Searches
  - Models
  - Conclusions

---

**MICDE**
MICHIGAN INSTITUTE FOR COMPUTATIONAL DISCOVERY & ENGINEERING
UNIVERSITY OF MICHIGAN

**MIDAS**
MICHIGAN INSTITUTE FOR DATA SCIENCE
UNIVERSITY OF MICHIGAN
Overview of today’s session

Graduate Certificate in Computational Discovery and Engineering

PhD Program in Scientific Computing

Graduate Certificate in Data Science

MS students: Multidisciplinary Design Program

mdp.engin.umich.edu
UM Graduate Computational Science Training
www.MICDE.umich.edu

Ken Powell, Director of PhD in Scientific Computing
Krishna Garikipati, Director MICDE
Mariana Carrasco-Teja, Assistant Director MICDE
The development and innovative use of mathematical/computational algorithms/models for research, science and engineering, data analysis and interpretation, product development, and forecasting

**Computational Science** is now widely accepted as the third pillar of science, complementing theory and experimentation

U-M offers several opportunities for specialization in this booming field, and engaging with other researchers in computational science
Computational Science Research

Acoustics  Blood flow  Materials Science

Computer Architecture  Physics  Current Smoking Prevalence

Hydrology  Space and Climate  Public Health
Computational Science Research

- Computational chemistry, biology
- Computational statistics, operations research, complex systems
- Computational business (analytics), finance, econometrics
- Computational social sciences (automated information extraction systems, social network analysis, social geographic information systems (GIS), complexity modeling, and social simulation models)
- ...
MICDE’s Mission

MICDE, created in 2013, is an initiative of the University of Michigan College of Engineering, College of Literature, Science and the Arts, and Advanced Research Computing

Our mission is to advance new paradigms of computational science, which cut across application domains, to enable the discoveries that will define society for the next decade and beyond.
MICDE today

Two Educational Programs:
- Ph.D. in Scientific Computing
- Graduate Certificate in CDE

141 Affiliated Faculty
31 Departments
8 Schools/Colleges

Research Centers:
- Storage-Enabled Comp Sci.
- Data-driven Comp Phys.
- Software Infrastructure

Seminar series, Symposia, Faculty Workshops

Outreach and Industrial Engagement
Educational Programs

- PhD program:
  - Offers opportunity for much deeper specialization in computational science

- Certificate program:
  - Open to all M.S. and Ph.D. students
  - Lightweight – recognition of exposure to / knowledge of field of computational science
PHD IN SCIENTIFIC COMPUTING

Ken Powell, Director
Prof. Aerospace Engineering
Ph.D. in Scientific Computing

- Must be pursuing a PhD in a home department at U-M
- Thesis topic and committee composition must reflect an emphasis on scientific computing
- 18 credit hours
  - 3 courses (9 credits) in numerical methods and
  - 3 courses (9 credits) in computer science and computing applications outside home department
- One of the prelim questions must be related to scientific computing
- Meeting the requirements appends “and Scientific Computing” to their diploma (e.g. PhD in Aerospace Engineering and Scientific Computing)
Ph.D. in Scientific Computing

- Established in 1989
- Prior to 2015:
  - 85 degrees granted; ~15 students enrolled at any given time
Ph.D. in Scientific Computing

Currently:

- 81 students enrolled [22 Depts., 6 schools/colleges]
Ph.D. in Scientific Computing

- 108 alumni (23 on or after 2015)
Application Procedures:

- Talk to your academic advisor about your interest.
- Send an email to Prof. Powell and Bonnie Bryant, program administrator at micde-phdapp@umich.edu to set up a meeting and finalize your planned courses.

Students are encouraged to apply to the program after having completed one term of doctoral work, but prior to being promoted to candidacy status.
RACKHAM GRADUATE CERTIFICATE IN COMPUTATIONAL DISCOVERY & ENGINEERING

Krishna Garikipati, Director Prof. Mechanical Engineering & Mathematics
Graduate Certificate in CDE

- Nine graduate credit-hours (3 courses) (methodology and application)
  - All courses with substantial computational content are allowed
  - If a course is not listed in our site, please ask us about it
  - One course can be double counted with other degree

- Attendance at MICDE Annual Symposium and MICDE Seminar Series
  - At least 7 seminars

- CDE-related non-credit experience
  - E.g., internship, research, another course, MDP
Graduate Certificate in CDE

- Established in 2013; 48 alumni
- 41 students currently enrolled [18 Depts., 4 schools/colleges]
Graduate Certificate in CDE

- To enroll
  - Talk with your advisor
  - Follow the procedures described at https://micde.umich.edu/certificate/application-procedures/
MICDE Student Resources

- MICDE organized and taught courses:
  - Methods and Practices in Scientific Computing
  - Data-Driven Modeling of Complex Systems
- $4K$ top-off fellowships
- CAEN accounts for non-engineering students
- Outreach opportunities
Scientific Computing Student Club

- ~80 members
- Active throughout the year:
  - Alumni Series: Eric Harper (NRL) Nov 9, 2018
  - Flux Tour October 11, 2018*
  - 2018 Machine Learning Workshop
    Tuesdays @ 6:00 pm starting October 4, 2018
  - Happy hour @ ABC Brewery October 18, 2018

More information
micde.umich.edu/sc2
Questions about the certificate in CDE or the PhD in Scientific Computing?

Go to micde.umich.edu/academic-programs or email micde-contact@umich.edu
## Characteristics of Big Data (Biomed)

<table>
<thead>
<tr>
<th>BD Dimensions</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td>Harvesting and management of vast amounts of data</td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
<td>Wranglers for dealing with heterogeneous data</td>
</tr>
<tr>
<td><strong>Incongruency</strong></td>
<td>Tools for data harmonization and aggregation</td>
</tr>
<tr>
<td><strong>Multi-source</strong></td>
<td>Transfer and joint modeling of disparate elements</td>
</tr>
<tr>
<td><strong>Multi-scale</strong></td>
<td>Macro to meso to micro scale observations</td>
</tr>
<tr>
<td><strong>Incomplete</strong></td>
<td>Reliable management of missing data</td>
</tr>
</tbody>
</table>

**Example:** analyzing observational data of 1,000’s Parkinson’s disease patients based on 10,000’s signature biomarkers derived from multi-source imaging, genetics, clinical, physiologic, phenomics and demographic data elements.

Software developments, student training, service platforms and methodological advances associated with the Big Data Discovery Science all present existing opportunities for learners, educators, researchers, practitioners and policy makers.
Graduate Data Science Certificate

Overview

The overarching goal of the Graduate Data Science Certificate Program is to train a cadre of skillful data scientists with significant multidisciplinary knowledge, broad analytical skills and agile technological abilities. The program emphasizes the practice of modeling using modern technology to handle large, incongruent, and heterogeneous collections of data. The Graduate Certificate for Data Science is approved by the Rackham School for Graduate Studies.

The program provides interactive data-centered training and involves 9 credits of courses and 3 credits of experiential training that require a written report on data analytics. MIDAS faculty from different disciplines provide mentorship and advising and the Institute offers merit-based top-off scholarships for graduate students enrolled in the certificate program. The Graduate Data Science Certificate Program is open for enrollment on a rolling basis. U-M graduate students from any field are eligible to enroll. Merit-based top-off fellowships may be provided. Minority and underrepresented students are strongly encouraged to enroll and complete the program.

http://midas.umich.edu/certificate
Graduate Data Science Certificate

- Open to ALL registered UM graduate students
- Course Requirements:
  - 9 graduate credits - Algorithms & Applications (AA), Data Management (DM), Analysis Methods (AM)
  - 3+ practicum credits – approved Data Science-related experience
- Attendance:
  - MIDAS Annual Data Science Symposium
  - Weekly MIDAS Colloquial Series (1 year)

http://midas.umich.edu/certificate
Graduate Data Science Certificate

Core Proficiencies

- Algorithms & Applications: core data science principles, assumptions, applications
- Data Management: basic protocols for data management, processing, computation, information extraction & visualization
- Analysis Methods: hands-on experience, modeling tools and analytics, real project setting

http://midas.umich.edu/certificate
Graduate Data Science Certificate
Prerequisites

<table>
<thead>
<tr>
<th>Prerequisites</th>
<th>Skills</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed Undergraduate Degree</td>
<td>Quantitative training and coding skills as described below</td>
<td>The DS certificate is a graduate program requiring a minimum level of quantitative skill</td>
</tr>
<tr>
<td>Quantitative Training</td>
<td>Undergraduate calculus, linear algebra and introduction to probability and statistics</td>
<td>These are the entry level skills required for most upper-level undergraduate and graduate courses in the program</td>
</tr>
<tr>
<td>Coding Experience</td>
<td>Exposure to software development or programming on the job or in the classroom</td>
<td>Most DS practitioners need substantial experience with Java, C/C++, HTML5, Python, PHP, SQL/DB</td>
</tr>
<tr>
<td>Motivation</td>
<td>Significant interest and motivation to pursue quantitative data analytic applications</td>
<td>Dedication for prolonged and sustained immersion into hands-on and methodological research</td>
</tr>
</tbody>
</table>

http://midas.umich.edu/certificate
# Graduate Data Science Certificate Competencies

<table>
<thead>
<tr>
<th>Areas &amp; Applications</th>
<th>Competency</th>
<th>Expectation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools</td>
<td>Algorithms</td>
<td>Working knowledge of basic software tools (command-line, GUI based, or web-services)</td>
<td>Familiarity with statistical programming languages, e.g., R or SciKit/Python, and database querying languages, e.g., SQL or NoSQL</td>
</tr>
<tr>
<td></td>
<td>Application Domain</td>
<td>Data analysis experience from at least one application area, either through coursework, internship, research project, etc.</td>
<td>Applied domain examples include: computational social sciences, health sciences, business and marketing, learning sciences, transportation sciences, engineering and physical sciences</td>
</tr>
</tbody>
</table>

Best practices for scientific and application programming, efficient implementation of matrix linear algebra and graphics, elementary notions of computational complexity, user-friendly interfaces, string matching
<table>
<thead>
<tr>
<th>Areas</th>
<th>Competency</th>
<th>Expectation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management</td>
<td>Data validation &amp; Visualization</td>
<td>Curation, Exploratory Data Analysis (EDA) and visualization</td>
<td>Data provenance, validation, visualization - histograms, QQ plots, scatterplots (ggplot, Dashboard, D3.js)</td>
</tr>
<tr>
<td></td>
<td>Data Wrangling</td>
<td>Skills for data normalization, data cleaning, data aggregation, and data harmonization/registration</td>
<td>Data imperfections include missing values, inconsistent string formatting ('2016-01-01' vs. '01/01/2016', PC/Mac/Lynux time vs. timestamps, structured vs. unstructured data)</td>
</tr>
<tr>
<td></td>
<td>Data Infrastructure</td>
<td>Handling databases, web-services, Hadoop, multi-source data</td>
<td>Data structures, SOAP protocols, ontologies, XML, JSON, streaming</td>
</tr>
</tbody>
</table>

http://midas.umich.edu/certificate
# Graduate Data Science Certificate Competencies

<table>
<thead>
<tr>
<th>Areas</th>
<th>Competency</th>
<th>Expectation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical Inference</strong></td>
<td></td>
<td>Basic understanding of bias and variance, principles of (non)parametric statistical inference, and (linear) modeling</td>
<td>Biological variability vs. technological noise, parametric (likelihood) vs non-parametric (rank order statistics) procedures, point vs. interval estimation, hypothesis testing, regression</td>
</tr>
<tr>
<td><strong>Analysis Methods</strong></td>
<td><strong>Study design &amp; diagnostics</strong></td>
<td>Design of experiments, power calculations and sample sizing, strength of evidence, p-values, FDR</td>
<td>Multistage testing, variance normalizing transforms, histogram equalization, goodness-of-fit tests, model overfitting, model reduction</td>
</tr>
<tr>
<td><strong>Machine Learning</strong></td>
<td></td>
<td>Dimensionality reduction, k-nearest neighbors, random forests, AdaBoost, kernelization, SVM, unsupervised learning. Transfer learning, active learning, ensemble methods, CNN</td>
<td>Empirical risk minimization. Supervised, semi-supervised, and reinforcement learning, multiview learning, instance learning</td>
</tr>
</tbody>
</table>

[http://midas.umich.edu/certificate](http://midas.umich.edu/certificate)
Graduate Data Science Certificate

Contacts

- Prof. Ivo Dinov, Director: dinov@umich.edu
- Kristen Thornton, Graduate Coordinator: thorntok@umich.edu
- Moira Dowling, Project Manager: mdowling@umich.edu
- Alison Martin, Administrative Assistant: aalison@umich.edu
- MIDAS information: midas-contact@umich.edu
- www.MIDAS.umich.edu

http://midas.umich.edu/certificate
MIDAS & MICDE

Student Resources

- ARC-TS
  - HPC Cluster, Data Science Clusters, storage
- CSCAR
  - Free consultations
  - Free workshops
- Seminar series and symposia
- Women in HPC chapter
- Student Clubs:
  - Scientific Computing Student Club
  - Michigan Data Science Team
You may enroll at any time but...

- **Certificates**
  - To be considered for enrollment in Fall, students will need to apply by August 1\(^{st}\)
  - To be considered for enrollment in Winter, students will need to apply by December 1\(^{st}\)

- **PhD**
  - Should consider the time it takes to finish all requirements
2019 MDP Cohort

Recruiting First Year through Professional Masters Students
Who We Are

• Engineering program that is open to every Undergraduate and Masters Student
  • **Apply what you learn** in class to engineering design challenges.
  • Gain the **technical and professional** skills necessary to thrive in engineering research or industrial settings, and
  • **Experience how multiple disciplines** collaborate within a team.
What We Offer

Research Teams

- Ongoing Research Teams in Faculty Labs
- Multiple Semesters (minimum 2)
- Student Leadership Positions

1-Year Industry Projects

- Two Semesters: complete Design Cycle
- Corporate, Government or Non-Profit Client
- Small Teams with dedicated Faculty Mentors
Who are the Students?

School of Music, Theater and Dance
Stamps School of Art & Design
School of Information
Ross School of Business
College of Engineering
College of Literature, Science & Arts
Taubman College of Architecture
School of Kinesiology...and more!
Kellogg Eye Center
2017 Team

1 CE
1 EE (graduate)
2 Computer Science
1 BME (graduate)
1 Biology

Bill Breznau

Stryker Project ‘16
Now: Sr. Engineer at Stryker in Mechanical Design
E-Nable
Non-Profit Team Designing 3D Printable Prosthetics for Children in Low-Resource Countries

1 EE (Graduate)
3 Computer Science
2 Biomedical Engineering
1 S.I. UI/UX (Graduate)
**Cadence for the Year**

*Industry Teams*

- Literature, Patent Review
- Define Customer Requirements
- Concept Development & Selection
- Test Methodology Development & Validation
- Design Expo & Final Delivery

- Weekly Dashboards
- 2 hours per week w/Faculty Mentor
- 1 hour per week w/Sponsor Mentor
- 8-10 hours per week of progress on the project
# How does it fit in my schedule?

## 1-Year Industry Teams

<table>
<thead>
<tr>
<th></th>
<th>WN 19</th>
<th>FA 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduates</td>
<td>ENGR x55 (3 credits)</td>
<td>BYO! (4 credits)</td>
</tr>
<tr>
<td>Masters Students</td>
<td>ENGR 599 (1-2 credits)</td>
<td>BYO! (2-3 credits)</td>
</tr>
</tbody>
</table>

## Faculty Research Teams

<table>
<thead>
<tr>
<th></th>
<th>WN 19</th>
<th>FA 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduates</td>
<td>ENGR x55 (2 credits)</td>
<td>ENGR x55 (2 credits)</td>
</tr>
<tr>
<td>Masters Students</td>
<td>ENGR 599 (2 credit)</td>
<td>ENGR 599 (2 credits)</td>
</tr>
</tbody>
</table>
What happens in the Summer?

Summers with MDP

Competitive Funding Selection: up to $5k Summer Research Stipends. Apply in February. Limited number available, not guaranteed.

Most Industry Project sponsors offer Guaranteed Internship Interviews, some guarantee Internships.
So what are the teams?

Browse, Sort, Filter all 46 Teams & Projects online

http://mdp.engin.umich.edu
Gameday VR

Students on this team will develop a first-of-its-kind application for existing off-the-shelf Mixed Reality headsets from Microsoft. The application will allow individuals (such as patients at CS Mott Children's Hospital) who cannot attend a sporting or live event in person to 'virtually' experience the excitement of a sporting event via augmented/mixed reality. Students will leverage best practices in contextual design to understand the needs and desired experience of their end user.

Likely Majors
- CSE/CS-LSA
- SMID-PAT, SI
- Screen Arts & Cultures
- A&D
- IOE, PSYCH
- Any

Meet the Gameday VR Team
Preview Night:
Oct. 2nd, 6 - 8 pm
Project Fair:
Oct. 3rd, 12 - 4 pm
Credit Towards Degree Available

Project Areas
- Programming
- Video and Environment Capture
- Database Design
- U/I/UX

Likely Majors
- CSE/CS-LSA
- ROB, MIDAS
- Data Science
- ECE/ROB
- ME, MSE
- Applied Physics
- IOE, Math
- AUTO, GAME

Meet the Bosch Team
Preview Night:
Oct. 2nd, 6 - 8 pm
Project Fair:
Oct. 3rd, 12 - 4 pm
Credit Towards Degree Available

Apply Online at:
mdp.engin.umich.edu
by Oct. 14th

Robust Path Planning for Automated Vehicles

Autonomous vehicles require an intelligent system to sense and process information, not only from objects and obstacles, but also from the road itself. Students on this team will develop a confidence system for autonomous vehicles to identify free paths or "drivable surfaces" to guarantee the integrity of maneuvers, such as automated lane changes on the highway.
**Optimization of Current Control in Electric Motors**

Interior Permanent Magnet machines are a type of motor that has a rotor embedded with permanent magnets. Compared with the SPM (surface permanent magnet), this type of motor can reduce the risk of a magnet being peeled off by centrifugal force, and takes advantage of reluctance torque. IPM motors are widely used in the hybrid electric vehicles (HEV) sector. Students on this team will develop and validate a Simulink model to demonstrate effective power regulation to control the machine's torque, saving precious electric energy and reducing gas emissions.

**Autonomous Controls Development for Prototype Electric Vehicle Truck**

Developing autonomous capabilities in vehicles is critical for improving safety, overall societal efficiency and promoting innovation. Students will develop the autonomous navigation system on Isuzu's electric vehicle truck and create a virtual testing environment to evaluate the robustness of its current navigation algorithms.

**Project Areas**
- Perception
- Computer Vision
- Controls
- Data Science
- Virtual Modeling

**Likely Majors**
- EE
- CS
- Robotics
- ME, CE
- Data Science
- STATS

**Meet the GM Team**
Preview Night: Oct. 2nd, 6 - 8 pm
Project Fair: Oct. 3rd, 12 - 4 pm
Credit Towards Degree Available

**Meet the Isuzu Team**
Preview Night: Oct. 2nd, 6 - 8 pm
Project Fair: Oct. 3rd, 12 - 4 pm
Credit Towards Degree Available

Apply Online at: mdp.engin.umich.edu by Oct. 14th
Secure Cloud Manufacturing

This research will make large-scale manufacturing systems safer, more secure, and more productive, enabling them to produce high-quality products for consumers at lower cost.

Sub-Teams
- Controls Subteam
- Cloud and Enterprise Subteam
- Simulation Subteam
- Apprentice Researcher

Likely Majors
- CSE/CS-LSA
- CE, ME, EE
- SI, MIDAS
- MATH, IOE
- STATS, Any

Meet the Researchers
- Preview Night: Oct. 2nd, 6 - 8 pm
- Project Fair: Oct. 3rd, 12 - 4 pm

Meet the Northrop Grumman Team
- Preview Night: Oct. 2nd, 6 - 8 pm
- Project Fair: Oct. 3rd, 12 - 4 pm

Opioid Addiction and Treatment Analytics

The student team will utilize a range of predictive data analytic techniques, including machine learning, to investigate opioid policy questions relating to opioid addiction and treatment.

Likely Majors
- Data Science
- IOE, Stats
- Health informatics
- MIDAS
- CS, EE
- Public Health
- Nursing
- Biomedical Engineering

Apply Online at:
- mdp.engin.umich.edu
- by Oct. 14th

Credit Towards Degree Available
SOCR Big Data & Predictive Analytics

The Statistics Online Computational Resource (SOCR) is an online platform including educational materials, web-services, and advanced methods and tools in probability, statistics, and machine learning in the health sector. This team will develop an enhanced analysis and visualization toolbox with an emphasis on "Big Data" - very large datasets that are difficult to analyze and interpret in meaningful ways with basic probability and statistical methods.

Our Team Has
27 - 42 students

Sub-Teams
Programming (SOCRAT)
Analytics
Data Science Fundamentals
Apprentice Researcher

Likely Majors
CSE/CS-LSA
SI, STATS
Biostatistics
Bioinformatics
MATH, ENGIN
PHYS
(Math or Engin background)

Meet the Researchers
Preview Night:
Oct. 2nd, 6 - 8 pm
Project Fair:
Oct. 3rd, 12 - 4 pm

Apply Online at:
mdp.engin.umich.edu
by Oct. 14th

Credit Towards Degree Available
How to Apply

Attend either

October 2\textsuperscript{nd}: Preview Night 6 – 8pm in BBB
or

October 3\textsuperscript{rd}: Project Fair 12 – 4pm in DUDE

Meet the sponsors, bring a resume, shake hands, ask questions.

APPLY ONLINE BY 11:59pm on OCTOBER 14
Timeline

Attend a Recruitment Event (10/2 or 10/3)
Apply by 10/14
Offer Letters sent in November
Course Override for WN19 sent in December
Work begins in January!

mdp.engin.umich.edu
Questions?

joyadams@umich.edu

mdp.engin.umich.edu
**What are my choices as a Master’s student?**
Only the Certificates are available.

**What if the courses I’d like to count towards the certificate or degree are not listed on your website?**
Contact us, we likely can accommodate you.

**As a PhD student, can I do both the Certificate and the Joint degree?**
No.

**How do I choose which to do?**
The commitment for the Joint degree is greater: you need to do a prelim question on scientific computing, there are more courses required, and you need to take computer science. Choose which fits for you.