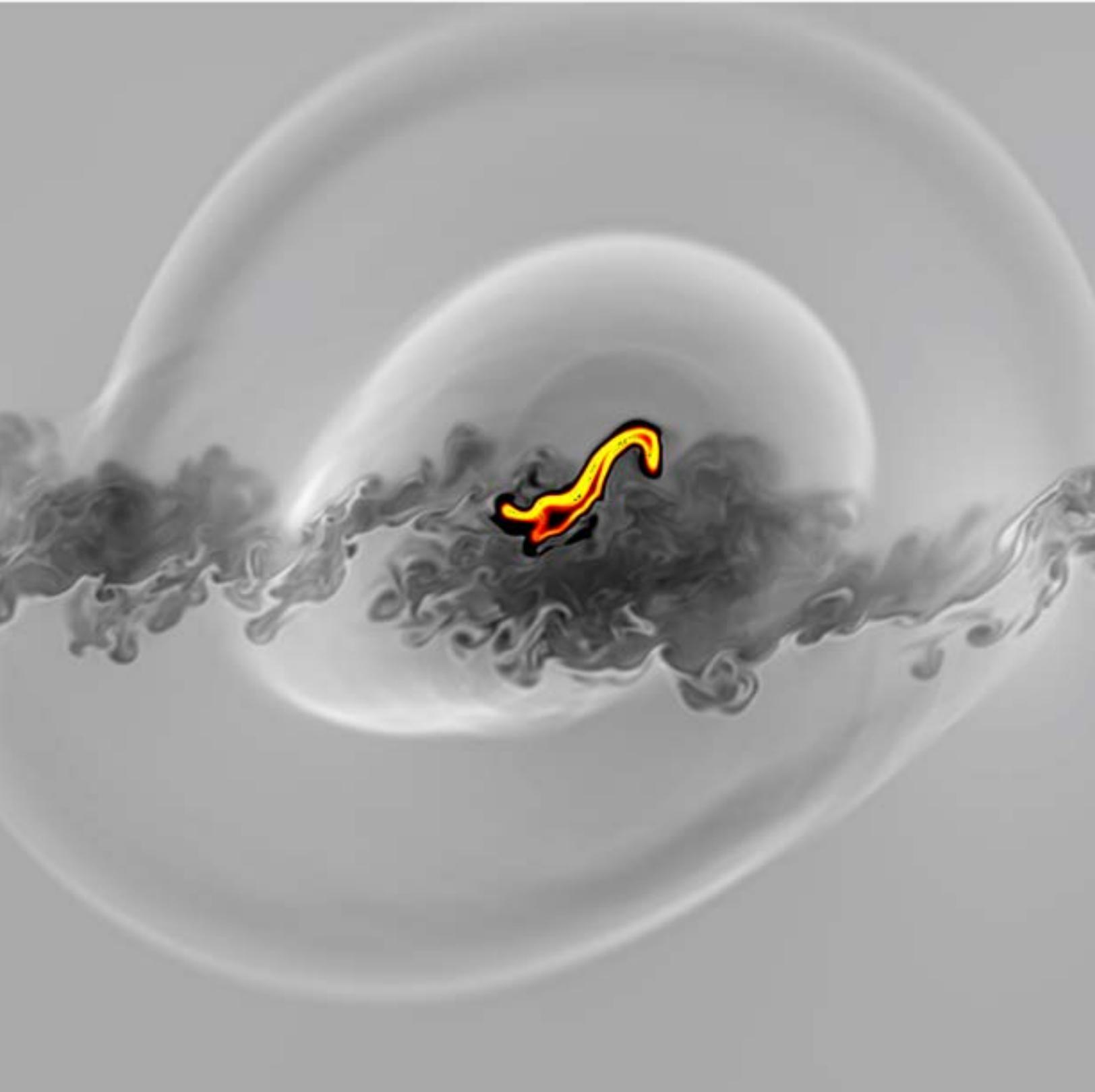


**Year in Review | FY 2018-2019**



**Advanced Research Computing (ARC)** leads advances in data-intensive and computational science at University of Michigan and beyond by providing U-M researchers a comprehensive and robust computing ecosystem.



**Eric Michielssen**  
Associate Vice President for Advanced  
Research Computing  
Professor of Electrical Engineering  
and Computer Science

**Sharon Broude Geva**  
Director, Advanced Research  
Computing



The **Michigan Institute for Data Science (MIDAS)** empowers innovative, data-driven research, leveraging U-M faculty expertise in areas such as transportation, learning analytics, health, social science, and music.



**Brian Athey**  
**MIDAS Co-Director**  
Professor of Computational  
Medicine & Bioinformatics



**Alfred Hero**  
**MIDAS Co-Director**  
Professor of Electrical Engineering  
and Computer Science



The **Michigan Institute for Computational Discovery and Engineering (MICDE)** advances new approaches to computationally intensive modeling and simulations across disciplines of science and engineering.

**Krishna Garikipati, MICDE Director**, Professor of Mechanical Engineering & Math

**Advanced Research Computing - Technology Services (ARC-TS)** develops, builds and maintains cutting-edge computing infrastructure resources available to the entire U-M research community.

**Brock Palen, ARC-TS Director**



**Consulting for Statistics, Computing and Analytics Research (CSCAR)** provides workshops, training, and consulting services to help U-M researchers from all disciplines take full advantage of the University's advanced computing resources.

**Kerby Shedden, CSCAR Director**, Professor of Statistics

Cover illustration: A simulation from Assistant Professor Jesse Capecelatro, Mechanical Engineering, show heat release from combustion.

# A message from Eric Michielssen, Associate Vice President for Advanced Research Computing

Advanced Research Computing at U-M (ARC) continues to grow in order to provide a comprehensive ecosystem of resources and support for all researchers on campus engaged in computational or data-intensive science.

The breadth and depth of research at the University of Michigan that depends on computational and data science is staggering. ARC's challenge is to provide this growing group of investigators all the tools they need to get the job done. That means anything from access to high performance computing resources to consulting on research methods to assistance with funding submissions — and much more.

The results have been impressive, as researchers assisted by ARC or its units have published groundbreaking findings and garnered significant funding support:

- Rada Mihalcea, professor of Electrical Engineering and Computer Science, developed an algorithm that's better than humans at detecting fake news
- Rafael Meza, associate professor of Epidemiology, is a principal investigator in a new, \$18 million NSF-supported tobacco research center
- Students and faculty from the Michigan Data

Science Team are helping the city of Flint deal with its water crisis

- Shawn McKee from the Department of Physics is part of a new NSF-funded research consortium on high energy physics software.



To help promote new research achievements going forward, this year ARC has added new resources in a number of areas:

- Consulting for HPC code optimization and data science
- The new HPC cluster "Great Lakes" is nearing completion
- Data science services such as database hosting, data ingestion tools, and secure data enclaves
- A new Big Data cluster in partnership with Cavium

As an increasing number of researchers take advantage of computational and data intensive tools, ARC stands ready to provide the resources they need for the next generation of scientific discovery.

M   ADVANCED RESEARCH COMPUTING			
COMPUTATIONAL SCIENCE	DATA SCIENCE	TECHNOLOGY SERVICES	CONSULTING SERVICES
MICDE	MIDAS	ARC-TS	CSCAR

Advanced Research Computing (ARC) at the University of Michigan consists of the Michigan Institute for Data Science (MIDAS), the Michigan Institute for Computational Discovery and Engineering (MICDE), Advanced Research Computing - Technology Services (ARC-TS), and Consulting for Statistics, Computing and Analytics Research (CSCAR).

# Research Highlights

## **New tobacco research center will use advanced data analytics**

The University of Michigan School of Public Health will house a new, multi-institutional center focusing on modeling and predicting the impact of tobacco regulation, funded with an \$18 million federal grant from the National Institutes of Health and the Food and Drug Administration.

The Center for the Assessment of the Public Health Impact of Tobacco Regulations will be part of the NIH and FDA's Tobacco Centers of Regulatory Science. MICDE will support the center's Data Analysis and Dissemination core by collecting national and regional survey data, conducting analysis of the use of tobacco products including vaping and e-cigarettes, and disseminate the resulting tobacco modeling parameters to other research centers and the Food and Drug Administration.

The center is led by MICDE affiliated faculty member Rafael Meza, associate professor of Epidemiology, and David Levy, professor of Oncology at Georgetown University.

## **Creating the Next Generation of Software for High Energy Physics**

The University of Michigan and MICDE are part of an NSF-supported 17-university coalition dedicated to creating next-generation computing power to support high-energy physics research.

Led by Princeton University, the Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP) will focus on developing software and expertise to enable a new era of discovery at the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland.

Shawn McKee, Research Scientist in the U-M Department of Physics, is a co-PI of the institute. His work will focus on integrating and extending the Open Storage Grid networking activities with similar efforts at the LHC.

## **Data-driven solutions for the Flint water crisis**

The Michigan Data Science Team (MDST) and its faculty advisors Eric Schwartz and Jacob Abernethy have been helping citizens of Flint, Mich., predict where lead-contaminated pipes may be as the city of nearly 100,000 struggles with its years-long water safety crisis. Using data from digitized versions of decades-old city records in combination with more recently produced datasets, the team's predictive models could save the city as much as \$10 million as it replaces pipes containing lead.

A paper describing the research, co-authored by Abernethy, Schwartz, Alex Chojnacki, Arya Farahi and Jared Webb, titled "Active Remediation: The Search for Lead Pipes in Flint," won the Best Student Paper award at the KDD Conference in London in August, 2018. DOI: 10.1145/3219819.3219896

## **Fake news detector works better than humans**

An algorithm-based system that identifies linguistic cues in fake news stories could provide news aggregator and social media sites like Google News with a new weapon in the fight against misinformation. Rada Micalcea, Ph.D., Professor of Electrical Engineering and Computer Science, developed the system and has demonstrated that it's comparable to and sometimes better than humans at correctly identifying fake news stories. The research was supported by MIDAS and the NSF, and presented to the 27th International Conference on Computational Linguistics. URL: <https://arxiv.org/pdf/1708.07104.pdf>

## ARC grants support U-M researchers across campus

Grants awarded through MICDE and MIDAS support a wide range of data-intensive and computational science.

**MICDE Catalyst Grants** promote innovative projects in computational science based on novelty, likelihood of success, potential for external funding, and potential to leverage U-M's existing computing resources. In Spring 2018, the Institute awarded funding to seven projects:

- Exploring Quantum Embedding Methods for Quantum Computing. Researchers: Emanuel Gull, Physics; Dominika Zgid, Chemistry
- Teaching autonomous soft machines to swim. Researchers: Silas Alben, Mathematics; Robert Deegan, Physics, Alex Gorodetsky, Aerospace Engineering
- Urban Flood Modeling at “Human Action” Scale: Harnessing the Power of Reduced-Order Approaches and Uncertainty Quantification. Researchers: Valeriy Ivanov, Civil and Environmental Engineering; Nick Katopodes, Civil and Environmental Engineering; Khachik Sargsyan, Sandia National Labs
- Deciphering the meaning of human brain rhythms using novel algorithms and massive, rare datasets. Researcher: Omar Ahmed, Psychology, Neuroscience and Biomedical Engineering
- Deep Learning for Phylogenetic Inference. Researcher: Jianzhi Zhang, Ecology and Evolutionary Biology; Yuanfang Guan, Computational Medicine and Bioinformatics
- Embedded Machine Learning Systems To Sense and Understand Pollinator Behavior. Researchers: Robert Dick, Electrical Engineering and Computer Science; Fernanda Valdovinos Ecology and Evolutionary Biology, Center for Complex Systems; Paul Glaum, Ecology and Evolutionary Biology
- Advancing the Computational Frontiers of Solution-Adaptive, Scale-Aware Climate Models. Researchers: Christiane Jablonowski, Climate and Space Sciences and Engineering; Hans Johansen, Lawrence Berkeley National Lab

**MIDAS Research Hubs** invest in cross-cutting research that takes advantage of emerging data science methodologies and applications. These research hubs leverage these investments by fostering collaboration between faculty and students in a variety of disciplines focused on some of the most important problems in our society, including Transportation, Health Science, Learning Analytics, and Social Science. In 2018, MIDAS awarded four grants, creating a new research hub — Music: Connecting Data Science and Music:

- *Understanding and Mining Patterns of Audience Engagement and Creative Collaboration in Large-Scale Crowdsourced Music Performances*  
Investigators: Danai Koutra and Walter Lasecki, both assistant professors of computer science and engineering
- *Understanding How the Brain Processes Music Through the Bach Trio Sonatas*  
Investigators: Daniel Forger, professor of mathematics and computational medicine and bioinformatics; James Kibbie, professor and chair of organ and university organist
- *The Sound of Text*  
Investigators: Rada Mihalcea, professor of electrical engineering and computer science; Anil Çamcı, assistant professor of performing arts technology
- *A Computational Study of Patterned Melodic Structures Across Musical Cultures*  
Investigators: Somangshu Mukherji, assistant professor of music theory; Xuanlong Nguyen, associate professor of statistics

For more, visit [midas.umich.edu/music](https://midas.umich.edu/music)

# Consulting & Training

## Consulting group expands HPC and data science curriculum

Consulting for Statistics, Computing and Analytics Research (CSCAR) continues to expand its workshop and consulting offerings to meet the needs of the increasing numbers of researchers on campus using computational or data-intensive tools in their research. CSCAR provides workshops and consulting in the following areas, among others:

### Statistics and Data Science

- Data analysis
- Basic and advanced statistical methods
- Statistical software support
- Geospatial analysis including Geographic Information Systems
- Design, analysis, interpretation and communication for research studies involving data
- Large-scale data processing and management
- Efficient computation with data
- Scraping, aggregating and integrating data



### Research Computation

- HPC code optimization
- Adapting scientific codes to parallel environments
- Making efficient use of HPC resources
- Android apps for research

# Infrastructure

## New HPC cluster “Great Lakes” nears completion

Advanced Research Computing - Technology Services (ARC-TS) is building a new high performance computing cluster called Great Lakes, which will replace the existing Flux cluster. Great Lakes represents a nearly \$5 million investment in HPC, and will offer researchers increased performance with more cores, faster cores, faster memory, and a more balanced network.

The cluster will be built with:

- Dell EMC PowerEdge C6420 compute nodes, PowerEdge R640 high memory nodes, and PowerEdge R740 GPU nodes
- Mellanox HDR 200Gb/s InfiniBand ConnectX-6 adapters, Quantum switches and LinkX cables, and InfiniBand gateway platforms
- DDN Gridscaler 14KX and 100 TB of usable IME (Infinite Memory Engine) memory

U-M’s HPC resources serve more than 2,500 active users at U-M for research ranging from aerospace engineering to molecular dynamics modeling to machine learning and artificial intelligence.

## University of Michigan becomes one of the first chapters of Women in High Performance Computing

In the summer of 2018, the University of Michigan was recognized as one of the first Chapters in the new Women in High Performance Computing (WHPC) Pilot Program.

“The WHPC Chapter Pilot will enable us to reach an ever-increasing community of women, provide these women with the networks that we recognize are essential for them excelling in their career, and retaining them in the workforce,” says Dr. Sharon Broude Geva, WHPC’s Director of Chapters and Director of Advanced Research Computing (ARC) at U-M “At the same time, we envisage that the new Chapters will be able to tailor their activities to the needs of their local community, as we know that there is no ‘one size fits all’ solution to diversity.”

“At WHPC we are delighted to be accepting the University of Michigan as a Chapter under the pilot program, and working with them to build a sustainable solution to diversifying the international HPC landscape” said Dr. Toni Collis, Chair and co-founder of WHPC, and Chief Business Development Officer at Appentra Solutions.

Women in High Performance Computing (WHPC) was created with the vision to encourage women to participate in the HPC community by providing fellowship, education, and support to women and the organizations that employ them. Through collaboration and networking, WHPC strives to bring together women in HPC and technical computing while encouraging women to engage in outreach activities and improve the visibility of inspirational role models.

WHPC is stewarded by EPCC at the University of Edinburgh.

## Student groups and certificate programs focused on data-intensive and computational science continue growth

ARC’s interdisciplinary institutes, MIDAS and MICDE, support several educational programs and student groups, participation in which continues to grow. The student groups have more than 500 active participants; combined enrollment in the graduate certificate programs is approaching 200.

### Student Groups

- Scientific Computing Student Club
- Michigan Data Science Team
- Michigan Student Artificial Intelligence Lab
- Computational Social Science Workshop
- Statistics in the Community

### Graduate Programs

- PhD in Scientific Computing
- Graduate Certificate in Computational Discovery and Engineering
- Graduate Certificate in Data Science



# ARC by the numbers FY 2018-2019

- All 19 schools and colleges at U-M use ARC consulting, training, and computing resources
- 4,000,000+ jobs on the Flux HPC cluster
- 90,000,000+ CPU hours
- 3,000+ consulting hours
- 4,000+ users of the Flux HPC cluster
- 350+ faculty members affiliated with the MIDAS and MICDE research institutes
- 500+ combined members of ARC-supported student groups

[arc.umich.edu](http://arc.umich.edu)

Advanced Research Computing is a unit of the University of Michigan Office of Research. The University of Michigan is a non-discriminatory, Affirmative Action Employer.  
© 2018 by the Regents of the University of Michigan.

 ADVANCED RESEARCH COMPUTING (ARC)

COMPUTATIONAL SCIENCE

DATA SCIENCE

TECHNOLOGY SERVICES

CONSULTING SERVICES

MICDE

MIDAS

ARC-TS

CSCAR