# GOERGEN INSTITUTE FOR DATA SCIENCE

## THE NEXT GENERATION OF DATA SCIENTISTS

> "What I'm learning here is helping me address a real societal need—getting more underrepresented youth interested in a STEM education. To do this, I'm mining Twitter, learning natural processing languages, and developing new software and analytical skills."
>
> *—Ling "Kelly" He '17*

▶ The University of Rochester offers BA, BS, and MS degree programs in data science. These are in keeping with the University's long tradition of preparing our students well for success in the world, while capitalizing on our distinctive strengths and harnessing technology in ways that will inform a variety of fields.

## BS in Data Science Students

### Ling "Kelly" He '17

Kelly He is a Xerox Engineering Research Fellow at the Hajim School of Engineering & Applied Sciences double majoring in math and data science and minoring in Spanish. Her interest in statistics, data mining, and artificial intelligence fueled her decision to major in data science. In fact, she recently began working with Jiebo Luo, a professor of computer science, on a research project designed to promote STEM education to high school students and, potentially, those starting their college careers. By mining data from Twitter and using Python (a natural language processing programming language), she hopes to identify students who would be interested in a STEM education in college and others who could be peer influencers and mentors for them.

"It's a great project because it feels like we are helping to solve a real societal need with data science," says He. "I get to learn a lot about software, data mining, and analysis in the process. It's exciting."

She has had other unique experiences here, too. In the spring of 2015, she participated in a leadership program with Ernst & Young in New York City. By interacting with professionals in its technology department, she saw how data science can be applied in daily life and in various fields, such as banking and security. In the fall, she will represent the University at the prestigious Grace Hopper Conference for women in computing.

## Lauren Kemperman '17

Lauren Kemperman has statistics in her blood. Her grandfather was Johanne Kemperman, a professor of statistics at the University. "Mathematics and statistics come naturally to me, but even if they didn't, I would still pursue them because of how practical they are—especially when they're combined with computer sciences. That's what the data science curriculum does, and that's why I love it."

As a data science major with a minor in statistics, Kemperman is excited about the modern approach to the curriculum and how it combines her interests in computers, mathematics, and statistics. She is also intrigued by the potential applications of data science. "Data science—and mining for meaningful information—is so important in every field."

Kemperman and fellow data science major Jean Chakmakas '17 have just started working with Wendi Heinzelman, a professor of electrical and computer engineering and dean of graduate studies for Arts, Sciences & Engineering, on a School of Nursing research project. They are designing an Android app to help people dealing with asthma. The app will ask questions about a user's health and point them to local resources.

"I'll be able to use math and statistics to develop questions for the users, and we will develop algorithms based on the data we collect," she says. "This will ensure the app is relevant to the user."

## Ulrik Soderstrom '16

Ulrik Soderstrom started out as an Earth and environmental sciences (EES) major and then switched to chemical engineering, but that still did not feel quite right. He started taking computer science classes and then heard about the new data science major. "I got passionate about it fast," he says. "All fields lead to data science."

Soderstrom is interested in using data science tools to help the world become more sustainable. That is why he is coupling the data science major with an EES minor. "I like concrete concepts, and I want to apply them to the renewable energy field," Soderstrom says. "I'm reaching out to solar companies now for an internship because I want to analyze energy market data, visualize energy grids, and understand how people's energy choices affect the environment."

Soderstrom is currently working in Cynthia Ebinger's lab as a research assistant. He is developing algorithms using Python and applying them to MATLAB to create 3-D models of volcanic activity in East Africa. This creates a picture of what is happening below the Earth's surface, which can help predict volcanic activity. "Having accurate models is essential to this work," he says. "Using data science is perhaps the only way to get there."

## Tyler Trine '16

Tyler Trine knew the data science major was for him as soon as he heard about it. And because he has always been interested in how the brain works, when he graduates next year with a degree in data science, it will be with a concentration in cognitive science.

"The data science major is the perfect marriage of my interests," Trine says. "It is a beautiful intersection of brain science, statistics, mathematics, and computer science. It takes a completely interdisciplinary approach, which is a huge Rochester strength."

Trine gained real-world experience when he was an intern with 1010data in New York City, a technology company focused on delivering business intelligence to just about every kind of company out there. When a company cannot handle the amount of data it collects, 1010data's powerful servers can, and they can provide actionable insight on the data that can improve marketing efforts and business performance.

The team there chose Trine because of a background rich in statistics and mathematics. "I'm confident that what I learned there, coupled with what I'm learning at the University, will create a really strong foundation for my career," he adds. After graduation, Trine plans to head straight into industry.

## MS in Data Science Students

### Andrew Straw '17 MS

For the last three years, Andy Straw has worked at the University of Rochester Medical Center in the Department of Biostatistics and Computational Biology. He helps researchers manage, integrate, analyze, and share data in support of multiyear, multimillion dollar grants. Prior to that, he spent five years building and using tools to help manage and make sense of biomedical data to support drug discovery at a pharmaceutical company.

Straw has a master's degree in computer and systems engineering, which included coursework in databases, software engineering, computer graphics, and artificial intelligence, but he has wanted to know more about statistics, machine learning, mathematical modeling, and parallel programming.

In his current position, he sees investigators struggling to make sense of large and diverse sets of experimental and clinical data. When the data science master's degree program was announced, he jumped at the opportunity, knowing it would give him the tools to help a broad range of users do more with their data.

"Statistical and mathematical models are so important to our understanding of how diseases, treatments, and the human body work," Straw says. "I want to be able to go back to my department and contribute to the development and application of these models to inform our research."

### Zach Taylor '15, '16 MS

Zach Taylor was a triple major at the University, where he earned degrees in economics, mathematics, and international relations. He intends to get a PhD in economics or political science and then will pursue a career in academia. Before he gets his PhD though, he will earn a master's degree in data science. Some have told him this is an unconventional step, but he knows the MS will benefit him.

## Careers in Data Science

- Advanced Technologist
- Analyst
- Analytics Consultant
- Associate Professor
- Business Information Consultant
- Chief Analytic Officer
- Chief Information Officer
- Chief Knowledge Officer
- Clinical Data Specialist
- Clinical Research Associate
- Collaborative Filtering Expert (Facebook, Netflix, Amazon, Pandora)
- Consumer Research Analyst
- Credit Risk Analyst
- Data Analyst
- Data Integrity Specialist
- Data Management Analyst
- Data Manager
- Data Mining Analyst
- Data Quality Analyst
- Data Scientist
- Enterprise Analyst
- Fraud Detection Analyst
- Government Surveillance Expert
- Health Informatics/ Information Management Director
- Health Informatics Specialist
- Health Information Manager
- Health Information Privacy Specialist
- Marketing Data Analyst
- Marketing List Analyst
- Model Validation Analyst
- Patient Information Coordinator
- Price Prediction Analyst
- Privacy/Security Officer
- Product Manager, Merchandising Metrics
- Project Manager
- Research Scientist
- Risk Analyst
- Scientific Programmer
- Software Developer
- Software Engineer
- Statistical Data Analyst
- Statistical Programmer
- Statistician
- Systems Engineer
- Technical Account Manager
- Technical Consultant
- Technical Expert
- Terminology Asset Manager
- Value Modeling Analyst
- Workflow Analyst

Taylor comments that economics and political science are fields that rely on highly quantitative methods. He also notes that data science is as applicable to these fields as it is to medicine and health care. "Social scientists today have a unique opportunity to harness the massive amounts of data available to tackle important questions," Taylor says.

The inherently interdisciplinary nature of the program appeals to Taylor. "It means I can take relevant courses in economics and political science," he says. "I will also take core classes to build and hone my computational skills. I know that exploiting big data to study politics will mean using advanced tools to approach problems in new and innovative ways. I'm fascinated by this program's potential to combine many different fields."

### Zidian Xie '16 MS

Zidian Xie is currently a postdoctoral researcher in the Department of Biostatistics and Computational Biology at the University of Rochester Medical Center. His work focuses on integrating biological databases for genetic testing. This requires biological and data analytic skills.

Xie is confident that a master's degree in data science will help him better identify and address important biological questions and conduct the most appropriate data analysis. He also recognizes how important it is for biologists and data scientists to communicate effectively with one another.

"With all the advances in medicine, we need people who can speak both languages," Xie says. "Right now, biologists don't necessarily have a thorough understanding of statistical concepts, and data scientists often find it challenging to communicate with them."

Over his 10-year professional career, Xie has seen how important it is to have a strong background in computer science, statistics, and mathematics. For instance, in a research position he held for four years at the University of Hawaii, he had to rely on other people to analyze his data

"Data science helps us to extract understanding from large-scale data. It provides a pipeline that transforms data to meaningful knowledge to action."

*—Henry Kautz*
*the Robin and Tim Wentworth Director*
*of the Goergen Institute for Data Science*



▲ Students analyze the data gathered by way of optical recording and imaging in the lab of director of the Center for Visual Science David Williams.

because he did not have the right skill set. "Data science will accelerate research in my field and in every field in the future," he adds. "Every student needs these skills."

### How You Can Help

Funding for professorships, research, graduate fellowships, and undergraduate scholarships will help the Goergen Institute for Data Science leverage existing strengths and help build relationships. Please consider a gift that supports these researchers and their work in data science.

For more information, contact
**Stephen A. Dare**
*Senior Associate Vice President*
*University Advancement*
(585) 275-7530
stephen.dare@rochester.edu

Learn, Discover, Heal, Create—And Make the World Ever Better
**www.rochester.edu/data-science**