

Delving into DATA SCIENCE

Students complete their first year in Rochester's new program for harnessing the power of digital information.

*By Kathleen McGarvey
Photographs by Adam Fenster*

When it comes to tackling social problems and fueling business, experts agree that data—the vast volume of data that the digital world generates—increasingly is indispensable, a resource that scholars, business leaders, and consumers have only begun to tap.

In response a new field has emerged, called data science. It's the creation and application of new methods for collecting, curating, analyzing, and making discoveries from large-scale data, in areas such as energy and the environment, economics and politics, health care, and marketing. The areas of application are only limited, really, by data scientists' imaginations.

This spring, the inaugural crop of students in Rochester's new data science programs—at the undergraduate and master's degree levels—are completing their first year of study at the Goergen Institute for Data Science, a program of the Schools of Arts & Sciences. Some are completing their degrees, having finished a year of graduate study or having gotten started early by taking prerequisites for an undergraduate degree. Others are just beginning. But what unites them all is an enthusiasm for discovering ways to make the most of information.

“We want to teach the students the fundamentals—computer science and statistics—and then let them explore some of the applications in different areas. That's what I think is unique,” says Henry Kautz '87 (PhD), the Robin and Tim Wentworth Director of the Goergen Institute for Data Science and a professor of computer science. He notes that programs emerging at other schools are often more specialized—dedicated to business analytics or health analytics, for example—but that the Rochester program provides a strong foundation for every area of focus.

Having a degree in data science “tells employers that these are people with a broad range of skills on the statistics side and the computer science side. They're highly in demand,” he says.

That's something to which Greg Munves '04 can attest. He's president and chief operating officer of 1010data, a New York-based company dedicated to big-data discovery and data sharing. “You need folks with Renaissance skills,” he says. “So instead of searching through people deep in computer science but with a little statistics or math, now you get people who have just the right amount of proficiency in each of those disciplines.”

Two of this year's graduates are headed to 1010data. “For the size company we are, we have a disproportionate number of U of R alumni involved,” says Munves. “With the advent of the data science program, it's become even more relevant for us to look to U of R for good talent.”

And once the program was launched last fall, students voted with their feet. Original predictions for the program's first year were for five graduate and 12 undergraduate students. Instead, year one brought 26 and 30 enrollees, respectively. And those numbers could double next year, says Kautz.

The long-term goal for the master's degree program is 50 students—the number for which the lab space was designed. For the undergraduates, there is no ceiling.

“We'll grow the program as much as the students demand,” Kautz says. 



Ling (Kelly) He '17

Degree: Bachelor's in data science

Data science concentration area:

Computer science, statistics, and mathematics

Hometown: Hangzhou, China

When Ling (Kelly) He '17 arrived at Rochester, she was thinking about majoring in economics or math. But she also became interested in the concept of “big data”—the huge data assets generated by digital life that can reveal important patterns. When she learned that the University was kicking off a new program, she wasted no time in signing up. Now she's a double major in data science and mathematics, with a minor in Spanish.

Last summer, He began doing research under Jiebo Luo, an associate professor of computer

science with expertise in data mining, machine learning, and other areas of investigation for data science. The project to which she contributed used data collected from social media to connect students who aren't interested in STEM—science, technology, engineering, and mathematics—with people who could nurture their curiosity.

“It's amazing how you can make use of resources to find a pattern, or do something for the social good or for other applications in real life,” He says.

She didn't know much about issues surrounding STEM education when she first joined the project, but now, she says, she's “more passionate about it” than she ever thought she would be. “I can see the power of data science because I know more about it.”

DATA-DRIVEN: “I can see the power of data science because I know more about it,” says He of her first year in the program, during which she contributed to research using social media data to connect students with mentors who can pique their curiosity about science, technology, engineering, and math.

This summer she has an internship in New York City with business-management consulting firm Ernst & Young, in a technology-advisory program of the company's financial services office.

“There's a lot to learn. It's a new field with a lot coming out really quickly. It's very fast-paced. Sometimes it can be overwhelming because we're still learning the basic techniques. It can be overwhelming—but that's what keeps you learning and going forward.”



Ian Manzi '18

Degree: Bachelor's in data science

Data science concentration area: Economics

Hometown: Kigali, Rwanda

Ian Manzi '18 has been intrigued by computers ever since he was a child. But it's not computers alone that spark his interest.

"A lot of people think of people interested in computer science as 'antisocial'—people and machines. But I am not that," he says. "I feed off the energy of being with people."

He's convinced he has found an ideal fit in data science for his analytical mind and gregarious personality.

Manzi arrived at Rochester unsure of what he wanted to pursue. He started with a major in computer science, but he found it too abstract. And so he moved this year to data science, which he says keeps the quantitative aspects of programming that appeal to him

but also offers more opportunities for practical problem-solving.

Last year he completed an internship as an operations manager in his hometown of Kigali, Rwanda, helping local schools to transition to digital record-keeping.

"Marketers and programmers don't speak the same language," he says. "I understand both, and I helped the project to grow."

It's the act of communication that seems to fascinate Manzi most. He says that he wants to create computer platforms that help people converse across disciplines. "The answers we need for societal problems may be on the 'other side'" of divisions, he says, and communication is essential.

Fostering conversations and looking for answers is something he knows a thing or two about. This spring, Manzi and Derrick Murekezi '19, a geology major from Nyagatare, Rwanda, received a \$10,000 grant from Davis

MAKING CONNECTIONS: Drawing together different fields and points of view is part of what attracts Manzi to data science. "I feed off the energy of being with people," he says.

Projects for Peace, a competitive national program that promotes peace and intercultural understanding. The two will use the funding to implement their summer project, "Critical Thinking for Peace: Sustaining Peace in Post-Conflict Regions," in Rwanda this July. They'll plan, set up, and operate a young leaders' peace camp for high school students, where camp-goers will reflect on Rwanda's 1994 genocide.

And while he didn't conceive of the project as having a data science dimension, he's beginning to see ways that he might potentially work that in.

One thing data science has cultivated in him, he says, is the "ability to relate to different fields."



Anthony Margetic '16

Degree: Bachelor's in data science

Data science area of concentration: Biology

Hometown: Syosset, New York

Tyler Trine '16

Degree: Bachelor's in data science

Data science area of concentration:

Brain and cognitive sciences

Hometown: Wolcott, New York

This summer, two of the data science program's first new graduates will be headed to Manhattan, to begin working at data discovery and sharing company 1010data.

Anthony Margetic '16 and Tyler Trine '16 got their feet wet at the company last summer, when they were interns there.

"It was exciting to be there for just two and a

half months," says Trine. "I want to go along for the ride."

They're already off to a roaring start, having devoted themselves to the program the year before it was even launched, structuring their schedules to get prerequisites out of the way before the program began. "It's been a jam-packed two years," says Trine.

But those two years have given them a new way of seeing the world. "Data science is a way of thinking," says Margetic. "You want to get to the questions that no one else is looking for."

It's the fluidity and interdisciplinarity of data science that first drew him. "I love learning about all kinds of stuff. I didn't want to be stuck in one field," he says.

For Trine, discovering the field of data science was finding an intellectual home for the kind of thinking he was already doing. He had started reading philosophy and picked up a book on semiotics, the study of how we make meaning.

NEW YORK, NEW YORK: With new degrees in hand, Trine (left) and Margetic are Manhattan-bound, to start work at data discovery and sharing company 1010data.

The linguistics minor was intrigued by "what it means to be informed," he says. "It became a preoccupation. I wanted to understand how people do it on a deep level, how we take in rich sources of information intuitively."

Says Trine, "It used to be that data was scarce, and the main enterprise was squeezing the most out of it. Now we're data-rich, but still information-poor." Data scientists like Margetic and Trine hope to change that.

Margetic says that having the phrase "data science" attached to a degree helps set him apart because it highlights a breadth of skills. "If you're a math person or an engineer, you have a subset of the skills I have as a data scientist."

Francesca Romano '16 (MS)

Degree: Master's in data science

Data science area of concentration:

Computation and statistical methods

Hometown: Rotterdam, New York

When Francesca Romano was an undergraduate at Siena College, she majored in math, with a minor in computer science.

"The thing I loved about math, the thing that drew me to it, was using my problem-solving skills and applying them to all kinds of problems," she says.

Data science seemed to her the best way to continue to do that. This spring, the Rotterdam, New York, native is earning her master's degree in the field.

"It's made me confident in my analytic and programming abilities. Everything that we're learning, we can apply to all sorts of different problems—health care, market research. It's a really diverse degree."

Although opportunities are plentiful, job hunting can actually be tricky, she says. While a lot of employers are realizing they need to hire people to help them make data-driven decisions, they're not yet experienced in how to describe what they're looking for. "I had to be creative in seeing what companies are calling these positions," she says.

Her creativity has paid off. She's accepted a position in cardiovascular services at Ellis Hospital in Schenectady, New York, where she'll be analyzing data from cardiovascular patients to identify trends or patterns that may be useful to the hospital in providing care.

"Society is becoming more and more tech-savvy and technologically advanced," she says. "We're collecting massive amounts of data, from phones, from everything we buy online, and more. Even if you're not totally aware of what data science is and how it applies to you, I think it's important to know that it's become an important part of business and decision making."

One thing Romano particularly likes about Rochester's program is the range of students. "I think it's been great that the first year they accepted people from so many different backgrounds," she says. "It's been really nice to see that everybody has a bit of a different strength—and that we can combine our strengths to accomplish something."

PROBLEM SOLVER: Romano says that data science draws on her problem-solving skills. "It's made me confident in my analytic and programming abilities. Everything that we're learning, we can apply to all sorts of different problems—health care, market research. It's a really diverse degree," she says.





Tiffany Sinclair '15, '16 (MS)

Degree: Master's in data science

Data science area of concentration:

Health care

Hometown: Evanston, Illinois

Wherever you turn, data science is there, says Tiffany Sinclair '15, who this spring is completing her master's degree in data science.

"It really is in every field," she says. "We've collected so much information over the decades, and now we finally have the ability to do something with it."

For Sinclair, that something is contributing to public health. Her goal is to become a data analyst in a health-related position, and she plans

to find opportunities to travel and examine how health care works in different countries.

The master's program concentrates on methods, she says, teaching students how to glean useful information from data and how to apply those skills in different areas.

Sinclair says she's always loved problem-solving and sorting through information to find meaning and draw conclusions. Studying data science has only sharpened those skills.

"It's a lot of fun—and challenging, not having come from a computer science background," says Sinclair, who as an undergraduate was a biology major and favored the pharmaceutical and analytical aspects of biology.

She says her favorite feature of the program

EARLY ADOPTER: Sinclair, who plans to pursue a data science career focusing on public health issues, says it's exciting to be in on the ground floor of an emergent field. "There's a sense that the field is really alive, and you can drive it forward," she says.

is that while the basic concepts stay the same, each class comes at them "from a different angle."

And it's exciting to be part of a brand-new field. "I have a sense of knowing so much—and of not knowing anything at all, because everything is still being studied," she says. "There's a sense that the field is really alive, and that you can drive it forward."