STRATEGIC PLAN

Destiny with Data

The University's investment in the burgeoning field of data science grows to $100 million.

By Leonor Sierra

THE UNIVERSITY IS PLANNING TO BE A BIG player when it comes to big data.

That was the message of an October announcement in which President Joel Seligman outlined a plan to commit $50 million—in addition to more than $50 million spent in recent years—to greatly expand the University's work in the burgeoning field of data science.

The commitment will include the creation of an Institute for Data Science, construction of a state-of-the-art building to house it, and as many as 20 new faculty members with expertise in the field.

"This is the top University priority for the University's 2013–18 strategic plans that were adopted by the Board of Trustees on October 11," Seligman said in announcing the plans during his opening remarks at the Rochester Big Data Forum 2013, at which researchers in data science from around the nation gathered for a day of interdisciplinary talks and discussions.

"Data science is a defining discipline of the 21st century," Seligman said. "By combining sophisticated analytic techniques with rapidly improving computational capabilities, data science can help extract useful information from the quintillions of bytes of data that are created every day. It is the foundation, for example, of data-informed, personalized medicine, is central to national security and defense, and has already changed online commerce."

Rob Clark, senior vice president for research and dean of the Hajim School of Engineering & Applied Sciences, says he expects that the new institute will have an impact on the Rochester region through collaborations with local companies and by helping start new companies.

"The investment opens up great

DATA BASE: A new building (above, in an early conception) will help leverage the work of Rochester researchers in big data, says Henry Kautz, chair of computer science.

opportunities for the translation to the commercial sector of the discoveries and new techniques that are developed, and we will also be producing highly trained specialists in this area," he says, pointing to the University's tradition of entrepreneurship. For example, he noted, faculty, staff, and alumni of the Institute of Optics have gone on to lead or found about 160 companies, many of them locally.

Home to the Health Science Center for Computational Innovation—which houses an IBM Blue Gene/Q supercomputer, one of the most powerful computers in the world—the University has ongoing research projects in data science–driven fields such as machine learning, artificial intelligence, and biostatistics. The new initiative is also expected to leverage existing collaborations with companies such as IBM and Xerox.

"Rochester researchers are already exploiting the tools of data science in their work," says Henry Kautz, chair of the computer science department and director of Rochester's big data initiative. "For example, data science has been a key part of research done here to model and predict the spread of infectious diseases, to track the

Science Cornerstone

The home of the new Institute for Data Science will frame a quadrangle of buildings housing science and engineering programs.
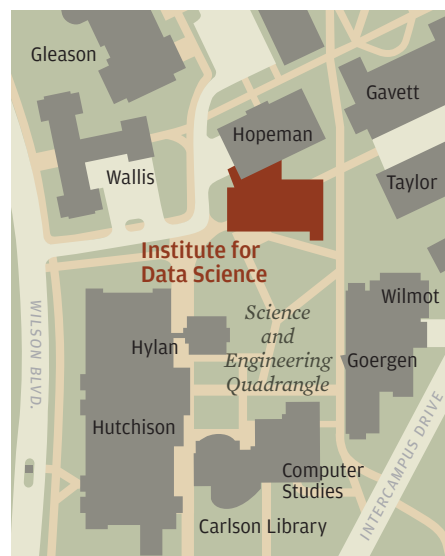
popularity of political ideas, to understand consumer preferences, and to predict the existence of planets."

Kautz says the University's expertise in

HOME COMPUTER: Rochester became one of the nation's five most powerful university-based supercomputing sites last fall when an IBM Blue Gene/Q supercomputer went online at the Health Sciences Center for Computational Innovation.

data science is dispersed across many departments and relies on individual groups of researchers to connect with each other to share their knowledge. The new institute will bring faculty members together with resources to enhance collaborations in data science in all fields, he says.

The new faculty members will be recruited in several departments: biostatistics, psychiatry, physics, computer science, political science, and others. But data science will be a critical component of their work, either as developers or users.

The additional faculty are also expected to open new areas of research as the institute develops. Three research domains have been identified as initial areas of focus: predictive health analysis, cognitive systems and artificial intelligence, and analyzing large-scale data as part of an effort known as "analytics on demand." **R**