How can Data Science Programs Contribute to the Research Environments of their Institutions?

Mujdat Cetin[1] and Ajay Anand[2]
[1]Robin and Tim Wentworth Director, [2]Deputy Director

Background on the Goergen Institute

Established in 2015, the Goergen Institute for Data Science (GIDS) and its associated programs are a top priority of the University of Rochester. Our data science program is designed to engage faculty, students, and the community in collaborations throughout the University, in the arts, sciences, engineering, business, education, medicine, and music. GIDS offers successful and sought after BA/BS and MS programs in data science that attract rapidly increasing numbers of candidates. 75 faculty members from around 30 departments across the campus, with shared interest in the theory, techniques, and applications of data science, are affiliated with GIDS. GIDS fosters industry engagement via two distinct entities that are part of its mandate: the Rochester Data Science Consortium and the Center of Excellence in Data Science.

The Institute has its origins beginning with computational initiatives funded through public/private partnerships dating back to 2005. The trajectory has been rapid, starting with the New York State/IBM/University of Rochester partnership that led to supercomputing capabilities. The Health Sciences Center for Computational Innovation was established in parallel and greatly accelerated our ability to leverage this infrastructure for the data science initiative. One of our accomplishments is the completion of Wegmans Hall, which now houses GIDS and was supported through the generosity of the Wegman Family Charitable Foundation.

GIDS builds on the University's existing strengths related to data science, many of which are dispersed across multiple departments and divisions. Major areas of expertise of GIDS faculty include:

- artificial intelligence and cognition;
- audio and music engineering;
- augmented and virtual reality;
- economics and business data analytics;
- healthcare analytics and digital health;
- human-system interfaces;
- imaging science, visual science, & optics;
- life sciences and biomedical data science;
- linguistics and social data sciences;
- materials design;
- multisensory human & machine perception;
- network data science; and
- robotics.

GIDS emphasizes industry engagement in its education and research activities. Undergraduate and graduate degree programs provide Capstone and Practicum project experience, respectively, through which students conduct real-world analytics projects using data provided by sponsoring organizations. GIDS provides the academic foundation for the Center of Excellence in Data Science (CoE) as well as the Rochester Data Science Consortium (RDSC). Created in 2017 and located in the heart of the Rochester Downtown Innovation Zone, the RDSC offers businesses and organizations a competitive edge through data expertise and data-driven solutions. RDSC staff apply university research to commercial problems, nurturing long-lasting relationships with local industrial partners. RDSC membership comprises major employers and universities in the Finger Lakes region, including L3Harris, Wegmans, Rochester Institute of Technology, and Corning, as well as new and growing startup companies. Through support from the

New York (Empire) State Development's Division of Science, Technology, and Innovation (NYSTAR), the CoE helps drive regional economic development by supporting research, training, and business development partnerships in data science. All of these efforts support the growth of data science as a major contributor to educational and economic opportunities in the area and beyond.

Fostering Interdisciplinary Research on Data Science across the Campus

Many data science institutes/centers/programs have an affiliation structure that is different from that of traditional departments. In particular, faculty often maintain their primary appointments with their departments while being affiliated with a data science program. While this non-traditional structure is essential and provides the opportunity to bring together researchers from a variety of disciplines, facilitating and maintaining meaningful sustained research interaction among the affiliated faculty is a challenge. Accordingly a fair question is: how can a data science program enhance the broad research environment of the University, producing research outcomes that cannot be achieved by traditional departments? We feel that two necessary steps data science programs should take towards that goal are:

(1) Maintain strong connections among affiliated faculty and researchers as a data science community on campus. Enable affiliated faculty to view the data science program as an integral dimension of their professional identities.
(2) Establish mechanisms and incentives to encourage and facilitate collaborative interdisciplinary research on data science. Make sure these mechanisms are sustainable and seamlessly adaptive.

With that perspective, in Fall 2019 we initiated an interactive process among GIDS affiliated faculty to make progress on both fronts. One concrete outcome has been the establishment of the GIDS Working Groups. GIDS Working Groups is aimed to serve as a mechanism to foster collaborative interdisciplinary research on data science and also as an activity-driven tool to shape and update the strategic priority areas of GIDS, effectively building on the strengths of the University of Rochester, while also enabling clusters of faculty to initiate new research directions to make an impact on emerging research areas collectively. A GIDS Working Group is composed of faculty from multiple departments, led by a chair or several co-chairs, with self-identified interest in an interdisciplinary research area related to data science. The following seven initial GIDS Working Groups represent the current research priorities and strengths of GIDS, and facilitate interdisciplinary collaborations on data science:

- Foundations of machine learning and artificial intelligence;
- Imaging, optics, and computer/human vision;
- Life sciences and biomedical data science;
- Health analytics and digital health;
- Human-data-system interfaces (including human-computer interaction, augmented and virtual reality (AR/VR), robotics);
- Complex systems and network data science;
- Economics and business data analytics.

We will present an overview of GIDS and its initiatives, and report on the activities of these working groups in our lightning talk. Over the long term, we will assess the effectiveness of this mechanism to contribute to the research environment of our institution.