

## Data Science Institute

### **Brief history of the institute**

Starting in 2014, a group of faculty from a diverse range of research areas began to network and develop a community of researchers who were interested in data science. The group held regular meetings to discuss how we might benefit from each other's expertise and to discuss the need for new curricula in data science. To support community-building activities, the group received some internal seed funding through a Vanderbilt initiative designed to support interdisciplinary collaborations. In 2017, Vanderbilt's Provost formalized this effort by appointing a university-wide working group to make concrete recommendations for new investments in data science. Following a year of careful strategic planning, the working group proposed the creation of a new trans-institutional research institute and new degree programs. The data science institute (DSI) at Vanderbilt was formally launched in 2018 and was granted financial support in the form of (i) a centrally located physical space, (ii) a commitment to hire a number of joint faculty, and (iii) startup funds to get the DSI up and running.

The DSI leadership team is truly interdisciplinary and includes faculty from Astrophysics, Biostatistics, Biomedical Informatics, Civil Engineering, Computer Science, Medicine, Political Science, and Psychology. The institute does not live within any of Vanderbilt's Schools or Colleges, but reports directly to the Office of the Provost.

### **DSI Mission**

- Promote data-intensive research and quantitative collaboration across Vanderbilt.
- Provide a forum to study the societal impact of 'big data'.
- Facilitate educational programs for the community at large, including a professional MS program in data science.
- Establish and strengthen new and existing connections with government, industry and non-profit entities who are active in the data science space.

### **Major accomplishments to-date**

- The DSI launched a two-year professional **Master's program in Data Science** with a brand new progressive curriculum that emphasizes practical workplace skills as well as foundational knowledge. The first class of 31 students started in fall of 2019, and the second class of 33 students just started this fall. As a result of a concerted effort to attract a diverse cohort of students, more than 50% of the new class consists of female students.
- Over the past two years the DSI has executed six successful **joint faculty searches**, together with the departments of Biomedical Engineering, Chemistry, Computer Science,

## Data Science Institute

and Political Science. These faculty are deeply engaged in the institute and they teach courses for the Data Science Master's program, as well as in their home departments.

- The DSI has hired a **team of professional data scientists** embedded within the institute. This team is engaged in a number of collaborative projects with researchers across campus and provides technical consulting and training to the community at large.
- The DSI has hired a first cohort of **postdoctoral fellows** who conduct research at the intersection of data science and fields such as anthropology, astrophysics, biomedical engineering, computer science, health policy, and political science.
- The DSI runs a **summer research program for undergraduates** who do projects in a variety of fields and also participate in a series of workshops to strengthen their technical skills.
- The DSI has fostered **industry partnerships** with several local and national companies to help educate students and connect them with internship opportunities.
- Vanderbilt is finalizing plans to launch an **undergraduate Minor degree in Data Science**, which will include several new courses and will be accessible to students from all backgrounds.

### Challenges and advice for new programs

- Building a truly interdisciplinary institute is difficult. Faculty are busy and have lots of demands on their time within their academic disciplines. It is critical to first develop a tight-knit community of faculty from different departments who believe in the mission and are willing to invest significant time in the effort.
- Hiring joint faculty with academic departments is important, but there are many risks involved. Search priorities of the department and the institute might not be aligned. Tenure criteria might not reward interdisciplinary work. There must be clear communication of expectations between all parties involved.
- Running an interdisciplinary degree program that lives outside of traditional academic departments is educationally valuable, but it poses several administrative challenges. The program might not benefit from existing support structures within Colleges and Schools. Borrowing faculty from other units to teach classes can be difficult.
- Attracting and retaining high quality data scientists is possible, but it is important to give them ownership of the institute mission and some degree of freedom in choosing what they work on.
- There is always a danger that a data science institute could become a 'service unit', since there is heavy demand in technical support across campus. Universities should establish separate mechanisms for programming, web, database, and research IT support.