



## **The Evolution to a School of Data Science**

Douglas C. Hague, Executive Director, School of Data Science, UNC Charlotte

### **Overview and History**

The University of North Carolina at Charlotte (UNC Charlotte) is an R2 university with over 24,000 undergraduates and nearly 6,000 graduate students. Data Science began at UNC Charlotte with the Data Science Initiative (DSI) that was originally conceived by a university wide-committee that formed in 2010. DSI launched a Health Informatics certificate in 2012. It was quickly followed by a Professional Science Masters (PSM) in Health Informatics (HI). These were followed by a certificate and PSM in Data Science and Business Analytics (DSBA) in 2013 [1]. These degrees were housed in the UNC Charlotte Graduate School and grew quickly to over 200 students. They were designed and continued to evolve through interaction and advice from local corporations that needed talent in these fields (large banks, health care entities, retailers, and other industries).

### **The Data Science Initiative**

The stakeholders at the university varied between the HI and the DSBA programs. While DSI was originally housed within the Graduate School, the HI degree and certificate operations were a partnership between the College of Health and Human Services (CHHS) and the College of Computing and Informatics (CCI). Similarly, the DSBA program was operated by the Belk College of Business (BCoB) and CCI. While funding for the program was primarily provided through School-Based Tuition Increment (SBTI), each change was run through committees in multiple colleges that decided independently from each other. This governance structure was difficult to maintain and it was not easy to move curricular changes through for approval due to a multipronged approval and debate process. Over the years, the Provost provided approximately 15 faculty positions to three different colleges for support of the DSI programs. Each College hired the faculty within their departments. There was a variation in the expectations for these positions across the colleges.

The SBTI funding supported for several primary functions, 1) staff to support students, 2) teaching assistantships, 3) part time faculty, 4) faculty stipends and course releases for leadership positions (e.g., graduate program directors) and 5) data infrastructure for faculty research. This last initiative purchased and maintained a large Hadoop based data storage facility. While originally targeted to be multi university, it never reached the vision and goals set out for it. Today the infrastructure hosts several data sets that require user-based security. The vision and mission as well as the overall architecture is currently under review.

Another focus of the Data Science Initiative that continues in SDS is to maintain strong industry connections. This resulted in two separate industry advisory boards (one for HI and one for DSBA). In addition, DSI launched, and SDS continues to host, an annual conference named Analytics Frontiers. This conference primarily consists of industry practitioner presentations and is hosted in Charlotte's uptown area close to many corporate offices. For the last several years and at industry request, we now host a separate Women in Data Science conference.

As DSI and its programs evolved (and HI became Health Informatics and Analytics, HIA), several growing pains emerged. First, the Graduate School was not set up to maintain a large focused program, so they passed the administration (but not governance) to one of the colleges (CCI). Second, the overhead associated with obtaining approvals from multiple colleges became increasingly difficult. Third, gaining instructional capacity from the various department chairs that were under their own capacity constraints set up conflicts. Finally, a decision was made to launch an undergraduate program that needed an academic home.

### **The School of Data Science**

For these and other reasons, the decision was made to change DSI to a School of Data Science (SDS). In December of 2019, UNC Charlotte Board of Trustees approved SDS. The academic unit of choice was a School structure vs department due to the interdisciplinarity that is both natural to data science and part of the vision for School. SDS governance expanded to include the College of Liberal Arts & Sciences (CLAS). This was a very intentional expansion to include not only the Mathematics and Statistics Department, but also all the social sciences and humanities departments.

Several key aspects of the founding of the School are: No additional internal or external funding was provided and, given COVID-19 impacts to state and university funding, the funding may be reduced from original expectations. Nearly all faculty are affiliate or joint appointments and participation is purely voluntary. As we launched the SDS faculty in August of 2020, there were more than 70 faculty affiliated with the School from 6 colleges and 19 departments. Governance is streamlined. The School is led by an Executive Director that is also a faculty member. The Executive Director reports and is governed by a Board of Directors. This Board of Directors is made up of the Deans of four colleges (CCI, CHHS, BCoB, and CLAS) and the Provost. SDS is developing a series of governance and curriculum committees for the School and no longer needs to go to each college for approvals.

A Bachelor of Science in Data Science is launching in the Fall of 2020. The requirements and pedagogy were carefully designed to meet several goals. First, the degree would focus on data science in the domain of social sciences and humanities. While domains like computer science, business and health are mentioned, the degree is built to include many more students than traditionally think of data science. Our program also chose to have no entry requirements above University admission. In addition, since UNC Charlotte has a large portion of transfer students, the degree was designed to be completed in two years by students with an Associate's degree. This specific constraint puts many limits and timing on the curriculum specified. Next, the data science courses are all taught through studio style pedagogy that has been shown to better enable the success of students from underrepresented groups [2]. Finally, we chose to integrate ethics and critical data studies directly into the studio courses. Each of the studio courses is being developed by a triumvirate of faculty. One each from computer science, social science, and philosophy. The blend of reading and writing critically with coding and statistics for a socially relevant challenge is expected to drive a unique learning experience for the students.

[1] B. Cukic, D. Hague and M. L. Maher, "An Innovative Interdisciplinary Undergraduate Data Science Program: Pathways and Experience," in accepted in IEEE Difference Makers Track of FIE 2020, October, 2020.

[2] S. Freeman and E. Theobald, "Is Lecturing Racist?," Inside Higher Ed, September 2, 2020.