

Title: Undergraduate Data Science Education in iSchools: Current Practices and Future Directions

Organizers

Loni Hagen, University of South Florida

James Andrews, University of South Florida

Hassan Zamir, Dominican University

Don Hamerly, Dominican University

Abstract

Since Song and Zhu (2016) concluded that undergraduate data science programs are “at the beginning step” in 2016 (Song & Zhu, 2016), many iSchools have initiated diverse data science programs to meet the shortage of data scientists. Educators of iSchools may have questions regarding how to define, and cultivate data science programs in iSchools, which originally have grown from the combination of computer science, statistics, and mathematics disciplines. Participants from institutions with undergraduate data science programs will present undergraduate data science recruitment processes, curriculums, barriers, and best practices through panel presentations. The session then has active discussion sessions with the audience. The session will conclude by initiating the development of a common repository of undergraduate data science curricula (syllabi, assessments, etc.), articulating a framework from which to build a successful undergraduate data science education model, and building consensus on possible future actions and venues for maintaining a network of undergraduate data science educators.

Purpose and Intended Audience:

The era of big data has ushered in a new class of problems that require multi-disciplinary efforts to tackle. Thus, the demand for data scientists have been increasing--IBM predicts demand for data scientists will increase 28% by 2020 (Columbus, 2017; IBM, 2017). Data scientists are expected to solve problems by extracting actionable information and insights from data, which implies a certain level of proficiency in technologies, programming, statistics, math, as well as communication, and business insights (Berry, 2018). In addition, data scientists should understand “the impact and application of data science” rather than functioning merely as number crunchers (Berry, 2018). The highly interdisciplinary nature of iSchools across these areas suggests they are positioned to meet these overwhelming requirements at the intersection of big data, people, and technology. Many iSchools already have undergraduate data science programs. Among the North American Members and Associate Members listed in the iSchools website, about 19 iSchools offer majors, minors, concentrations, or some electives related to data science, as of September 6, 2018. All these programs have unique flavors of educational models to best suit their unique bureaucratic structures, resources available, and characteristics of

student populations. It seem to be a good time to discuss undergraduate data science education in iConference to learn best practices, and to develop a framework for a successful undergraduate data science education in iSchools.

Intended participants include faculty (and students) at schools currently (or planning to) have undergraduate data science programs. Additional participants include industry practitioners who can articulate requirements for entry-level data scientists as well as future demands of skills.

The goals of this workshop are to produce three main outcomes. First, to create a community of data science education in iSchools. Second, to discuss curriculums, challenges, and best practices of educating undergraduate data science. Third, to develop a framework for successful and unique data science programs in iSchools. The major deliverable is a white paper summarizing the result of the session.

Proposed Format:

The first half of the session will be composed of a program overview and panel presentation with faculty from two iSchools: University of South Florida (Loni Hagen and James Andrew) and Dominican University (Hassan Zamir and Don Hamerly). Topics will be about recruiting students, how diverse students' backgrounds and education affect learning outcomes, curricula, interdisciplinary and interdepartmental collaboration, and job placement of former students. Q & A will be encouraged during the presentations. For the second half, participants and audience will have small group discussions to build a framework to provide a successful data science education in iSchool. Prior to the conference, we will send invitation emails and social media distribution of the SIE proposal to the 19 iSchools that provide undergraduate data science programs in order to solicit their participation in the SIE. In the solicitation, we will ask additional topics they want to discuss during the SIE. We will use **#iconf19_datascience** for all the social media communication.

Duration

90 min event.

Attendance

We expect to draw between 15 and 25 participants. Maximum 30.

Special Requirements

A room with presentation equipment. We prefer to have participants seated at smaller round tables to facilitate small group discussion.

References

Berry, A. (2018, June 18). Data scientists are in-demand and well paid - so why is there a skills gap? |

Computing. *Http://Www.Computing.Co.Uk*. Retrieved from

<https://www.computing.co.uk/ctg/opinion/3034263/data-scientists-are-in-demand-and-well-paid-so-why-is-there-a-skills-gap>

Columbus, L. (2017, May 13). IBM Predicts Demand For Data Scientists Will Soar 28% By 2020. Retrieved

September 5, 2018, from <https://www.forbes.com/sites/louiscolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/>

IBM. (2017, December 29). THE QUANT CRUNCH: HOW THE DEMAND FOR DATA SCIENCE SKILLS IS

DISRUPTING THE JOB MARKET. Retrieved September 5, 2018, from [https://www-](https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN&)

[01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN&](https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN&)

Song, I.-Y., & Zhu, Y. (2016). Big data and data science: what should we teach? *Expert Systems*, 33(4),

364–373. <https://doi.org/10.1111/exsy.12130>