

Information Science to Data Science: New Directions for iSchools

1. Organizing Committee	1
2. Abstract	1
3. Purpose and Intended Audience	2
3.1 Purpose	2
3.2 Audience	2
3.3 Rationale	2
4. Proposed Format	3
5. Relevance to the iConference	4
6. Duration	4
7. Attendance	4
8. Special Requirements	4

Workshop title: Information Science to Data Science: New Directions for iSchools

1. Organizing Committee

- Jane Greenberg, CCI, Drexel University (Co-chair)
- Xia Lin, CCI/Drexel University (Co-chair)
- Gobinda Chowdhury, Professor of Information Science, Chair, European iSchools, Department of Computer & Information Sciences, Northumbria University
- Sam Oh, Sungkyunkwan University: Library & Information Science and Data Science Department
- Virginia Ortiz-Repiso, Universidad Carlos III de Madrid
- Shigeo Sugimoto, University of Tsukuba: Graduate School of Library, Information and Media Studies
- Nanjing University in China (invited, representative will be confirmed)
- Wuhan University (to-be-confirmed)
- Penn State colleagues (invited, representative will be confirmed)

The organizing committee members are active participants in important national and international big data and cyberinfrastructure and data curation initiatives, and will engage colleagues from other iSchools. iSchool industry representative from companies such as Microsoft Research and global movements such as the Research Data Alliance will also be invited to participate.

2. Abstract

The iSchools represent the convergence of a range of disciplines (information/library science, computer science, communications, telecommunications, and information technology, while integrating aspects of social science, economics, and policy). The diversity of disciplines positions iSchools to contribute to data science research and workforce needs in both ordinary and novel ways. This workshop will provide a needed forum for international dialog on the iSchool

community's engagement with data science. The goals are to: 1) share iSchool developments in data science education and research, 2) articulate unique contributions and opportunities for iSchools in the data science, and 3) lay a foundation for continued international dialog among iSchools pursuing or seeking to pursue data science. A final report will summarize position papers, synthesize workshop discussions, outline action items for continued iSchool work in this area. An overall output, will be the foundation for a cohesive framework to enable the iSchool community dialog on data science.

3. Purpose and Intended Audience

3.1 Purpose

The purpose is to provide a needed forum for international dialog on the iSchool community's engagement with data science. An overall output, will be the foundation for a cohesive framework to enable iSchool community dialog on data science.

3.2 Audience

iSchool leaders, educators, and researchers pursuing and developing data science curricula, research agendas. Workshop organizers will also reach out to iSchool industry participants and seek engagement (e.g., Microsoft). We will also welcome iSchool student participation, and will solicit doctoral student participation, as the next generation of iSchool educators.

3.3 Rationale

Every day, we create 2.5 quintillion bytes of data— so much that 90% of the data in the world today has been created in the last two years alone.¹

The amount of data being generated is unprecedented, with workforce shortages by 2020 be predicted at 140,000 to 190,000 people with deep analytical skills, to higher numbers in the managerial and data analysts' area.² These numbers are sobering, and underscore both the opportunity and urgency for the iSchool community to lead data science education.

Along these lines, data science has rapidly evolved as a new academic area impacting curricula and research across many disciplines, including iSchools. This iSchool community has an opportunity to contribute to data science and offers unique blend of knowledge and skill, given information science interdisciplinary foundations in expertise in the collection, organization, manipulation, access, and use of information.

The iSchool community is significant in its focus on "relationships between information, people, and technology," and has an important role to play data science endeavors. In this context, it is not surprising that a number of iSchools have innovated data science curricula offerings, from courses and certificates, to fully endorsed degrees. Moreover, as the iSchool community expands, the number of iSchools pursuing data science is increasing. All of these developments have been swiftly executed to keep with the pace of big data and the digital data deluge, leaving little time for cohesive discussion. As this trend continues, it seems imperative that the iSchool

community launch a cross-institutional dialog on data science. This workshop provides a forum for this activity, by commencing an international dialog on the iSchool community's engagement with data science. The goals are to:

- 1) Share iSchool developments in data science education and research
- 2) Articulate unique contributions and opportunities for iSchools in the data science
- 3) Lay a foundation for continued international dialog among iSchools pursuing or seeking to pursue data science

4. Proposed Format

The workshop takes the form of position papers with breakout groups discussions. The format of this workshop will include an opening keynote to set the stage, followed by position papers, and breakout group for discussion.

The papers addressing the following three broad areas:

1. Current curricula - iSchool developments in data science education and research.
2. Opportunities - Unique contributions and opportunities for iSchools in the data science.
3. Framework for Continued Dialog - Designing a framework to support continued international dialog among iSchools pursuing or seeking to pursue data science.

The preliminary schedule is:

- 9:00-9:30 AM: Introductions and review of workshop goals.
- 9:30-10:00 AM: Keynote (Jeff Stanton, Syracuse, to.be.confirmed)
- 10:00-11:30 AM: Divide among 3 topics, position papers, and discussion (w/break as necessary)
- 11:30 AM-12:00 PM: Final framework discussion and collectively identify action items/next steps.
- Time permitting: Continue workshop for another half-hour of discussion.

A call for 2-page position papers (CfP) in the three broad area outlined above will be disseminated via the iSchool community and related information and library science, HCC, social networking, archives, and other community venues. The CfP will ask for position statements following a standardized format, and require authors to also submit discussion questions to guide breakout group discussion. The organizing committee will assess all submissions, and cluster and collate questions to guide breakout group discussion at the workshop. Position papers and the full workshop agenda will be made accessible prior to the workshop via the Metadata Research Center, College of Computing and Informatics, Drexel University (MRC/CCI/Drexel). The MRC/CCI/Drexel is leading collaborator with the global Research Data Alliance, NSF Regional Big Data Hubs, and global cyber-infrastructure, and all outputs will be made accessible via the MRC website.

The primary output will be a report that captures:

- Key insights detailing areas where iSchools currently provide a conduit for preparing successful data scientists.
- Educational and research-driven recommendations to iSchool for preparing graduates to address big data grand challenges and other ways for advancing data science.
- Actionable workshop outcomes and recommendations for continuing dialog and development of a cohesive community to advance iSchool contributions in data science.

5. Relevance to the iConference

The iSchool community has a growing interest in engaging in data science area as evident by curriculum changes and faculty hires. The format of the workshop (putting a broad callout for position papers) will enable those pursuing data science or seeking to engage in data science to participate. The workshop will also be open to individuals who have not submitted a position paper, but seek to engage in the dialog. Data science and related topics are keywords in the iSchool community, and there have been individual papers previously delivered on this theme, as well as discussion on this topic. This workshop seeks to take the discussion to the next level and engage the larger, growing, global iSchool community.

6. Duration

Half-day, following the iSchool 2017 requirement that all workshop must be ½ day.

7. Attendance

We anticipate approximately 40 participants. The workshop has an expansive design and can accommodate more attendees should registration exceed this number, as long as there is space.

8. Special Requirements

Basic technical requirements include set up for showing slides, and logistical requirements include seating for approximately 40 participants, or more if registration exceeds this number.

¹. IBM: What is big data? <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>.

². Manyika, J, et al. (May 2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute: <http://www.citeulike.org/group/18242/article/9341321>.