

Data Science Institute origins & plans

University Research Committee
Brian Yandell
7 March 2019

data science initiatives over the past 2 years

- Summer 2017-2019+: Data Science Hub (funded by UW2020 in May 2018)
- Fall 2017-2019+: Data Science Undergraduate Major Working Group
- Spring-Fall 2018: Computing Working Group*
- Spring-Summer 2018: RTAG Research Computing & Data White Paper
- Summer 2018: FoxConn Inst for Research in Science & Technology (in CoE)
- Fall 2018: Data Science (Institute) Planning Committee*
- Fall 2018-Spring 2019: CS/Stat/iSchool--WI in the Information Age* (in L&S)
- Spring 2019: VCRGE Data Science Institute creation*
- 2019: Research Computing & Data design* (with CIO)

*: charged directly or indirectly by Chancellor

bigger picture

- data science is connected to larger transformations of academy
 - data is reshaping our world
 - reorient UW to bring power of data science to all fields of study
 - future of data science will be shaped by insights from all other disciplines
- other academic activity
 - Moore-Sloan DS Environments
 - \$32M; see Creating Institutional Change (2017)
 - key sites: UWA eScience, UC Berkeley, NYU
 - others in MSDSE (15 including UMI MIDAS)
 - Data Science Leadership Summit Report (2018)
 - MIT's new \$1B College of Computing
- federal agencies
 - NSF 4 Big Data HUBs & 10 Big Ideas
 - Federal Big Data Research & Development Strategic Plan (2016)

Chancellor-charged ad hoc committees

- Data Science Planning Committee (8 Oct-29 Dec 2018)
 - Brian Yandell, Chair (Statistics, Horticulture)
 - Michael Ferris (Computer Science, WID), Katrina Forest (Bacteriology, RTAG chair), Martin Foys (English), Jan Greenberg (Social Work, OVCRGE), Young Mie Kim (Journalism), Lauren Michael (DS Hub), Robert Nowak (ECE), David Page (Biostatistics and Medical Informatics), Maureen Smith (Population Health Sciences), Alan Sorensen (Economics), Eric Wilcots (Astronomy, L&S Vice Dean) Stephen Wright (Computer Sciences)
- Data Science Institute Creation Committee (16 Jan-1 Mar 2019)
 - **Brian Yandell**, Co-chair (Statistics and Horticulture); Steven Ackerman, Co-chair (Atmospheric & Oceanic Sciences; Associate Vice Chancellor for Physical Sciences)
 - **Katrina Forest** (Bacteriology); Florence Hsia (History of Science; Associate Vice Chancellor for Arts & Humanities); **Robert Nowak** (Electrical & Computer Engineering); Petra Schroeder (Associate Vice Chancellor for Finance & Administration); **Maureen Smith** (Population Health Sciences); **Alan Sorensen** (Economics); **Stephen Wright** (Computer Sciences)

DSI overview

- **Data Science:** study, development, and application of methods that reveal new insights from data
 - Data: numbers, text, images, graphs, sounds, code, metadata, ...
- **Mission:**
 - perform cutting-edge *research* in the fundamentals of data science
 - *translate* this research by partnering with key application areas
 - *collaborate* with researchers across divisions to advance scientific discovery
- **Value Proposition**
 - leverage foundations of data science across all applications
 - build on UW-Madison's considerable strengths in multiple areas
 - collaborate with domain scientists to apply new algorithmic and computational tools
 - advance key application areas to enhance prospects for large-scale funding
- **Scope**
 - in: world-class data science research; broad partnerships & collaboration across campus
 - out: create / maintain research computing & data storage (RTAG white paper)
- **Components**
 - functions: research, translation, and collaboration
 - relationships & partnerships
 - organization & governance

organization & governance

- leadership and governance model
 - (Interim) Director reports to VCRGE
 - Assistant Director for Administration reports to Director
 - Leadership Team = Director, Assistant Director, Faculty Fellows
 - Advisory Board periodically will meet with Director
- organizational structure
 - Faculty Fellows (5-year with space in DSI) & Faculty Affiliates
 - Scientific staff
 - Data Science Facilitators operate Data Science Hub
 - Data Scientists (in DSI) & Research Scientists (with research teams)
 - Postdocs, graduate students (in DSI or affiliated)
 - Administrative staff
 - administration, communications, IT
 - Advisory board: internal spanning divisions; external advisors