DSI
1/ 16
D.E. Brown

Challenges
DSI
Research
Education

Data Science at the University of Virginia

Donald E. Brown, Director
Data Science Institute
University of Virginia
Charlottesville, VA USA
brown@virginia.edu

Society's most pressing challenges and nature's deepest mysteries frequently involve large, complex systems. The behavior of these systems is best understood by handling large amounts of data – thus the term "big data." To solve our most pressing challenges, we need to develop new ways to acquire, analyze, and make sense out of big data

Teresa Sullivan, President, University of Virginia

Societal Transformations

- Big Data is transforming marketing, retail, finance, manufacturing, energy production, health care, government, sports, and almost everything else
- Big Data is creating major new opportunities for cutting edge scholarship, research, and innovation
- There is a need for data scientists and literacy
  - McKinsey Global Intl:
    - 160,000 new workers with deep skills
    - 1.5m data-literate managers
  - Harvard Business Review: "Data scientist: the sexiest job of the 21st century"

Data Science Institute

A unique confluence of computation, science, engineering, mathematics, statistics, commerce, social science, humanities, law, & more

DSI
7/ 16
D.E. Brown

Challenges
DSI
Research
Education

The Creation of the DSI

- 2 years and 2 Big Data Summits
    - 170 attendees, 32 departments
    - Support from the UVA President
- Outcomes
    - Revealed data and analytic synergies among diverse disciplines
    - A new community emerged
    - Stimulated planning for collaborative, interdisciplinary projects, grants and an institute

DSI
8/ 16
D.E. Brown

Challenges
DSI
Research
Education

DSI Distinguishing Features

- Integrates all schools in a comprehensive university: arts and sciences, engineering and applied science, business, commerce, medicine, nursing, law, education, architecture, leadership, and continuing education
- Major endowment gift
- Advisory board of 14 members from 6 schools
- Advised by business and government partners
- Leverages differentiating research, scholarship and infrastructure
- Collaborates with data providers: Library of Congress, Census Research Data Center, National Radio Astronomy Observatory, Federal Agencies, Micron Center for Automata Computing, Multiple Health Centers

Range of Data Science Research

- Humanities
  - Examine theories of the mind through metaphor and word usage using "distance reading"
- Engineering
  - Data from finite element models of crash tests
- Environmental Science
  - Predictors of water usage and water quality
- Social Sciences
  - Maintained individual data distributed likelihood estimation

DSI
11/ 16

D.E. Brown

Challenges

DSI

Research

Education

- Systems biology
  - Links between environmental conditions and tissue response
- Healthcare
  - Health prognostics
- Astro-chemistry
  - How does chemistry emerge in the universe
- Sociology and Media Studies
  - Google and the Internet

Range of Research in Data Science

- Genomics
  - Genomic variation in somatic cells and disease
- Education
  - Data from video, body monitors, etc. to determine impact of instructional methods
- Commerce
  - Student analysis of web-based buying patterns

- M.S. in Data Science
- Minor in data science tailored to the student's major
- Internships with government and industry
- Continuing education courses and programs
- Jefferson scholar program for interdisciplinary graduate work

M.S. in Data Science

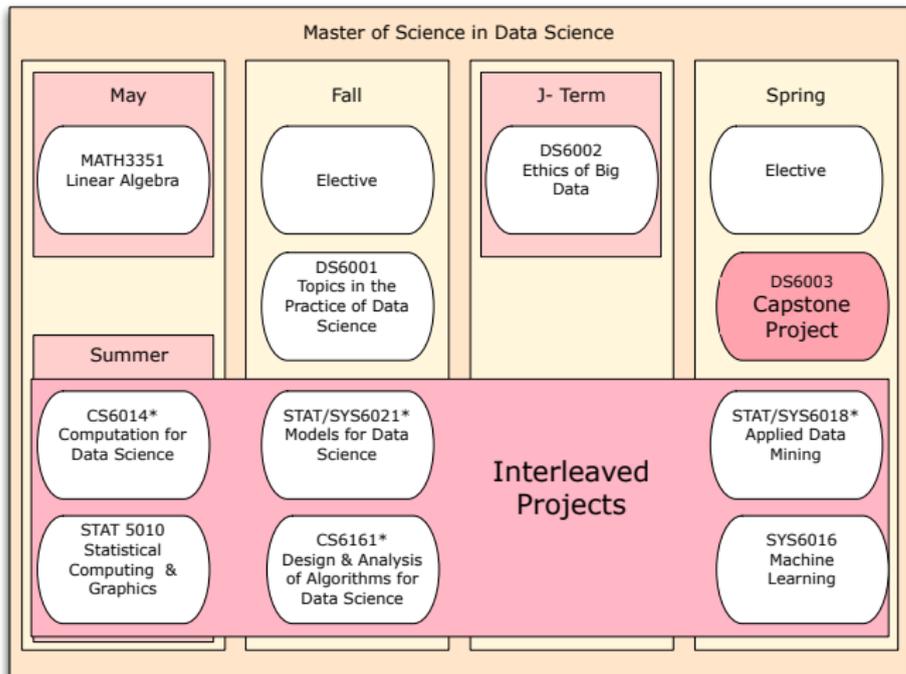# Distinguishing Elements of the M.S. in D.S.

- An integrated curriculum developed in consultation with practicing data scientists that builds progressively to a challenging capstone experience
- Applications and data drawn from a wide range of disciplines, e.g., science, business, and health.
- Practice course has: Distinguished guest lecturers, case studies in practice, & writing and presentations
- Course in data ethics, law, and policy
- 32 credit hours in 11 months, including a capstone course
- Job fair in the spring term