Data and Decision Sciences

Background: Data science has percolated into everyone's daily lives. From the social web and e-commerce (e.g., Amazon, Facebook, eHarmony) to smart thermostats (e.g., Nest) to crowdsourcing for better traffic data (e.g., Waze) to massive administrative sources (e.g., electronic health records), the data environment today is ubiquitous. Any fields not functioning in a data-driven manner are ripe for disruption.[1] In the era of Big Data, Virginia Tech believes every professional must become conversant in data science to be successful, and that the principled collection and analysis of data are vital guides for sense-making, discovery, and decision-making. We envision data science woven into the learning and research experience of students in every discipline. Virginia Tech graduates must be prepared to be leaders in a world where evidence-based decisions and systems are pervasive.

Current Virginia Tech Differentiators:
- Data science is being integrated into all disciplines: from the humanities and liberal arts to engineering, science, and business
- A "plus-data" concept takes advantage of existing collaborations to support degrees and programs that support the data revolution
- Virginia Tech has one-of-a-kind facilities, testbeds, and instrumentation featuring data science in action. These facilities support the "university as the sensor" viewpoint as a unifying theme
- The university is a research leader in all layers of the "data stack": new high performance data-intensive computing platforms, data mining/analytics algorithms, and visualization capabilities
- The university and its partners answer ambitious societal questions using data science and computational modeling in areas such as national security, epidemics, transportation, food security, international commerce, and social science.

Experience and Assets: Dozens of ongoing projects and examples of living laboratory programs illustrate the university's presence in the Data and Decision Sciences Destination Area.

Researchers in the not-so-distant past examined living systems one small slice at a time. We can now find complex connections ranging from activity at a cellular level to its effects on public policy. The Biocomplexity Institute of Virginia Tech is at the forefront of this scientific evolution, applying a contextual approach to answer some of the most pressing challenges to human health, habitat, and well-being. The Discovery Analytics Center brings together computer scientists, engineers, and statisticians to develop knowledge discovery systems in important areas of national interest. Its IARPA-sponsored EMBERS project scours vast quantities of public information in tweets, Facebook pages, news articles, blog posts, food prices, weather data, and economic indicators to forecast potentially disruptive societal events across the globe. The Center for Business Intelligence and Analytics undertakes quantitative and qualitative analysis of vast collections of business data to support planning and decision-making. Its work in mining online forums has led to earlier detection of automobile defects.

While data-pervasive environments are traditionally associated with STEM fields, Virginia Tech surmounts traditional boundaries by developing and enabling new technology for use of analytics in the

---

[1] Workplace Technology (A Special Report) Wall Street Journal, March 2016

social sciences and humanities. For instance, a "Digging into Data" project, supported by the National Endowment for the Humanities, capitalizes on expertise in computer science, history, and English to understand how newspapers shaped public opinion and represented authoritative knowledge during the deadly 1918 "Spanish" flu pandemic. In the humanities, gamification has brought to virtual life human tragedies concerning how we confront our greatest dilemmas.

Virginia Tech's versatility at training students in data-driven disciplines is apparent at all levels. Innovative new undergraduate majors (e.g., Computational Modeling and Data Analytics, Environmental Informatics), graduate degrees (e.g., Master of Science in Business Analytics, launching in 2016) and graduate certificates (e.g., Data Analytics, Urban Computing; launching in 2016) distinguish the university. Students are taught to work with community partners on how to gather information and use it ethically and professionally. In addition, our faculty, students, and partners also have access to data-collection tools and environments unlike any others on the planet. For example, Goodwin Hall, the world's most-instrumented building for vibration measurement, uses 241 accelerometers to give an unprecedented look at the secret lives of buildings: how weather and aging affect the structure and how its "metabolic" energy waxes and wanes, how the building can be a "first responder" to an emergency. Similarly, the hallmark laboratory of the Institute for Creativity, Arts, and Technology and the Center for the Arts is a highly adaptable space to create immersive audio and visual performances and conduct experiments. Here, researchers can simulate natural environments that people do not normally experience, e.g., a walk through a virtual tornado to understand weather and climate.

Examples of Targeted Hot Spots:

- **Data Miner in Your Pocket:** By 2047, Virginia Tech will develop technologies that will enable us to combine the power of data, analytics, and crowds so that every aspect of daily life is captured, harnessed, and mined to improve personal experience and forecast personal outcomes.
- **Data Journalism:** Data journalism will be the primary way by which results are communicated to the broader public. Data provides the medium by which democratization of science happens.
- **Internet of Analytics**: Virginia Tech's emphasis on "university as a sensor" will enable an emphasis in natural and built environments (e.g., future buildings, coupled water-food-energy systems, ecosystems support for human habitability, unmanned aerial vehicles/drones).
- **Precision Medicine:** Our emphasis on combining data science and mathematical modeling and expertise in human aspects enables us to significantly contribute to national "big data" challenges such as precision medicine (The White House's Precision Medicine Initiative has been touted the ultimate Big Data project.)
- **Community Learning Systems:** Virginia Tech is working with Washington, D.C., metropolitan area communities to liberate and repurpose their data flows to provide the data driven evidence necessary for communities to build an equitable and sustainable social transformation.
- **Knowing how people and organizations use information:** A reluctance or inability to act on information may influence how people make decisions, not lack of data. This research focuses on uncovering institutional problems in using information for problems such as head injuries in sports, the harms of tobacco, and health decisions.