Strengthening the University’s Capacity to Advance Knowledge and Serve the Public

Through Research, Scholarship, Creative Arts and Innovation

STRATEGY 4: Strategic Research

STRATEGY 5: Research Infrastructure and Services

What is the big idea?
Why do we need to do this?
Why does this make us distinctive?
How do we do it?
How long does it take?
How will we know we have succeeded?
Signature cross-Grounds institutes at UVa

• An institute for large scale, complex data analysis (Big Data)
• A unique confluence of computation, science, engineering, medicine, mathematics, statistics, commerce, social science, humanities, law, and more
What is Big Data?

- <2003 – 5B GB of information
- Today – 5B GB every 10 min (Social media)

Real time acquisition/analysis (markets)

Social media, images, EMR, genomics

When the volume, velocity or variety creates a problem

The challenges
- Management
- Computation
- Analytics
- Visualization
- Ethics
- Training

Examples
Commerce, NSA, elections, Boston marathon
What is the big idea?

Distinguishing examples of Big Data at UVa

• Chemistry of the Universe (Pate)
• Large scale, interdisciplinary, and interactive computational models (Learmonth)
• Cyber-physical systems, wireless sensors, and wireless health
  – Time series analysis of cardiac sensor data (Moorman)
• Computer-driven molecular engineering
  – Multi-drug resistance (Kasson)
• Systems modeling of complex regulation
  – Signaling (Janes)
  – Metabolism (Papin)
Why do we need to do this?

**It is transforming our society**

- 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
Why do we need to do this?

There is an opportunity

• Interviews of faculty, chairs, deans
• Computational modeling is
  – Present throughout the University - but unconnected
  – Outstanding in several areas
  – An area of development in the College (QC) and a focus in SEAS and SOM
  – Modeling and simulation are a Commonwealth priority
• Complex data sets (Big Data) are ubiquitous and untapped
Why do we need to do this?

The Big Data Summit revealed a need

• Format: short “meet and inform” talks
• Introduced by President Sullivan
• 170 people attended from 32 departments
• Summit outcomes
  – Revealed data and analytic synergies among diverse disciplines
    • DIA and dance (Wiesner)
  – Revealed our strengths and opportunities
  – A new community emerged
  – Stimulated planning for collaborative, interdisciplinary projects, grants and an institute
Why does this make us distinctive?

We have a comprehensive research university on a single campus

• This breadth is a distinguishing feature
• It provides the opportunity to offer unique curricula in data science and literacy
• It provides enormous potential for synergy
• It leverages differentiating research, scholarship and infrastructure across Grounds
Why does this make us distinctive?

UVa has differentiating research & scholarship impacted by Big Data

• Sciences
  – Chemistry of the universe
  – Environmental and ecosystems modeling
  – Computer-driven molecular engineering
• Humanities
  – Data-enabled humanities
  – Digital citizenship
    • Ethics and guidelines for production, collection, analysis, dissemination of big data
• Social sciences
  – Quantitative psychology
Differentiating research & scholarship impacted by Big Data (cont)

• Medical sciences and healthcare
  – Genomics of diabetes and genomic analysis
  – Wireless health
  – Systems modeling of cellular data

• Computation
  – Data integration
  – Large scale, interdisciplinary computational models (Bay Game)
  – Cyber-physical systems & wireless sensors

Unique opportunity for education and training for the new era of data driven activities
Why does this make us distinctive?

Differentiating infrastructure & enabling technologies for Big Data

• Computation
  – High performance and multi-institution grid computing
  – Cybersecurity

• Proximity to DC
  – Data-driven agencies
  – Intelligence community (NGIC & DIA)
  – Library of Congress (Culpepper) pipeline (in progress)
  – Census Research Data Center (in progress)

• GIS (geospatial information systems; Geospatial Resource Center)

• Advanced manufacturing connections

• NRAO (National Radio Astronomy Observatory)

• Micron Center for Automata Computing (in progress)
How do we do it?

In the wake of the Big Data Summit a new community emerged

• Big data prospectus
  – Presidential priority
  – Charge to coordinate and facilitate cross-school activities in Big Data
  – A team approach – VPR, Provost, CIO

• Planning committee
  – 14 faculty, 12 departments, 4 schools
  – Addressed infrastructure, services & facilities, curricula, community building, research areas and infrastructure of distinction and need
  – Input and dialog with stakeholders
  – Town hall meetings
  – Deans and University Library
  – Corporate and government partners

Strengthen the University’s Capacity to Advance Knowledge and Serve the Public
Through Research, Scholarship, Creative Arts and Innovation

PILLAR 2 | STRATEGY 4: Strategic Research
Develop transformative curricula

- Masters (& Certificate) in Data Science
  - Uniquely conjoins data acquisition, handling, analytics, computation, visualization and ethics
  - Multiple languages and software packages
  - Projects in multiple domains (transport, health, commerce, environment)
- Courses for non specialist students
- Undergraduate minors – content-related but domain-specific
- Broad-based undergraduate data literacy
  - 2 new Big Data courses (humanities & sciences)
  - Revise introductory math to include stat, computation, linear algebra, and calculus
  - Introduce Big Data topics broadly (Resource Center)
- Establish internships and partnerships (industry and government)
- Courses and certificates for continuing and executive education
How do we do it?

Identify resource needs and coordinate facilities and services

• Two levels
  – Routine
  – Research opportunities (collaboration)
• Coordinate resources, instruction, service
• Areas
  – Computation
  – Data management
  – Analytics
    • Partner with industry
  – Visualization
    • CS, Architecture, University Library
  – Ethics
    • New “Center for Data Ethics” (SEAS, CLAS, Law, SOM)
Build a new collaborative community

• Seminars and symposia
  – UVa Big Data Summits
  – Theme based workshops
    • Identify needs, synergies, and opportunities
  – Big Data symposium and seminars (planned)
  – Include external partners

• Jefferson Trust Team Building RFP
  – New student driven collaborations
    • Analysis of MOOCs (Sarah Turner (Curry,Econ), Dan Spitzner (Stat))
    • Data visualization (Shayne Pierce-Cottler (BME), Jeana Ripple (Architecture))
Identify partnership opportunities

• Institute and Directorship naming
• Named partnerships that share analytic software, cloud storage, and computing
  – The large spectrum of capstone and faculty project domains provide synergy and highly credible test and application sites
  – Software specific training for internships, jobs
• Named Centers of Excellence and Professorships in areas of distinction
• Student internships and fellowships
How long does it take?

Year 1

- **Director:** Appoint an inaugural director; initiate a search for a permanent director
- **Curricula:** Complete planning for a MS and Certificate in Data Science, undergraduate minor(s), courses for data literacy; initiate approval processes; disseminate the new curricula
- **Infrastructure:** Identify and prioritize needs; develop and coordinate plans for interdisciplinary training and services; implement new storage infrastructure and cores to accommodate big data; Census RDC online
- **Faculty:** Align unit-based hires that with Big Data needs and opportunities (statistics, visualization, machine learning); identify targets for distinguishing cluster hires
- **Research & scholarship:** Plan a major big data symposium; issue an RFP for student-initiated collaborative research
- **Partners:** Develop a fund-raising plan and identify partners for naming opportunities; identify and engage corporate and government partners for education, internships, and research
How long does it take?

Year 2

- **Director**: Identify and appoint a permanent director
- **Curricula**: Launch new graduate and undergraduate curricula; plan new elective courses that match student need; initiate courses in data literacy
- **Faculty**: Unit based hiring in areas of need that align with Big Data; initiate coordinated hiring in an area of strategic opportunity
- **Infrastructure**: Implement new storage system for multiple levels of performance, security, scale and backup; upgrade core network infrastructure to support multiple parallel virtual networks; seek funding for an advanced computational resource in Virginia; Library of Congress pipeline
- **Research & scholarship**: Catalyze and fund seed proposals for new research collaborations and directions; submit grant applications for collaborative research and teaching
- **Partners**: Engage research, resource, and teaching synergies with corporate and government partners; establish agreements for analytics, cloud storage, and computing

Strengthen the University’s Capacity to Advance Knowledge and Serve the Public Through Research, Scholarship, Creative Arts and Innovation

**PILLAR 2 | STRATEGY 4**: Strategic Research
How long does it take?

Years 3-5

• **Curricula**: New faculty develop courses in areas of programmatic need and student interest; reorganize the introductory math curriculum; establish a “resource center” that has big data examples for non major courses

• **Faculty**: Hires in areas of strategic importance

• **Infrastructure**: Upgrade infrastructure to accommodate new faculty and projects

• **Research & scholarship**: Develop centers of research excellence; biannual UVa Symposium on Big Data; grant applications based on new collaborations and faculty

• **Partners**: Establish partnerships and naming opportunities for emerging centers of excellence; implement distance learning and executive education

Strengthen the University’s Capacity to Advance Knowledge and Serve the Public

*Through Research, Scholarship, Creative Arts and Innovation*

**PILLAR 2 | STRATEGY 4**: Strategic Research
How will we know we have succeeded?

UVa research & teaching will push the agenda of “Top Jobs 21” & STEM development

- Research Centers of Excellence
  - Complex and interactive modeling across grounds
  - Wireless systems design and implementation
  - Cybersecurity
  - Genome analysis and variance in health and ecosystems
  - Text and image analytics across the humanities, arts, science, commerce, and law
  - A Center for Data Ethics
  - The Quantitative Collaborative - quantitative social science

- New data oriented curricula
  - Cross disciplinary masters and minor curricula & capstone projects
  - Undergraduate data literacy
  - New introductory mathematics curriculum

- A new institutional focus on collaboration and synergy
How will we know we have succeeded?

Interdisciplinary collaborations will be available for data management, analysis, and computation across Grounds

- **Humanities – Pasanek**
  - Examine theories of the mind through metaphor and word usage using “distance reading”

- **Engineering – Lach, Stankovic, Calhoun**
  - Data from low energy and wireless sensors embedded in people and structures

- **Education – Pianta, Dexter, Meyer, Bradshaw**
  - Data from video, body monitors, etc. to determine impact of instructional methods
  - MOOCS – student behavior (response times, answer changes, eye tracking) during test taking

- **Genomics – Hall, Rich, McConnell, Harrison**
  - Genomic variation in somatic cells and disease
  - EMR and personalized medicine

- **Healthcare – Heath**
  - Effectiveness of nurse practitioners under ACA
  - Effectiveness of nurse-physician student/resident teams on patient care

- **Commerce – Netemeyer**
  - Student analysis of web-based buying patterns

- New faculty will come here to engage in cutting edge research and teaching
How will we know we have succeeded?

Enhanced efficiency and coordination

- The Libraries
  - BioConnector (SOM library)
  - Data services (University Library)
- The Quantitative Collaborative
- UVACSE (engineering) and SHANTI (humanities)
- Census Research Data Center
- Library of Congress Pipeline
- Center for Data Ethics
How will we know we have succeeded?

- **Director:** Distinguished reputation, vision, leadership
- **Curricula:** Enrollment, job placement, student and faculty feedback, new revenue (MS & continuing Ed), enhanced student recruiting
- **Infrastructure:** Student and faculty use, revenue, feedback
- **Faculty:** Enhanced research reputation and recruiting, impact on research and scholarship, teaching effectiveness
- **Research & scholarship:** New collaborations, research directions, publications, interdisciplinary grants
- **Partners:** Engagement of corporate and government partners, named Professorships and Centers

Education, research, and scholarship at UVa will be transformed!

*Strengthen the University’s Capacity to Advance Knowledge and Serve the Public Through Research, Scholarship, Creative Arts and Innovation*

**PILLAR 2 | STRATEGY 4:** Strategic Research