

School of Data Science

Phase II Faculty Senate Submission

Prepared by the Data Science Institute

Last Modified April 17, 2019

Executive Summary	5
Background	6
I. Mission and Goals	7
Definition of data science	7
Mission	8
Goals	8
What is the intellectual purpose of the school?	11
What body of knowledge does the school encompass?	12
How does the proposed school further distinguish the University of Virginia?	12
What degrees are to be offered?	12
What will be the nature of the research contributions of the members of the school?	15
What services will be delivered to the Commonwealth of Virginia and the broader world?	16
II. Need for a School	18
Why a School and not an Institute?	18
Demand for the academic program/needs assessment/market analysis	20
Does the proposed school address a core need that UVA should meet to compare favorably with comparable universities?	22
Employment opportunities for the school's graduates	23
Review of competitive programs at comparable institutions	24
Results from preliminary vetting of the proposed school	26
Student target market	27
Impact on existing academic programs at the University of Virginia	27
AAC Requests for the SDS Phase II Document	27
Comments from three departments	35
Department of Statistics	35
Department of Mathematics	36
Department of Computer Science	36

III. Academic & Research Programs	38
 Founding Principles	38
Fit with UVA's mission	40
How will the new school fit within the context of existing academic programs?	40
 Research	40
 Capstone Projects	40
 Presidential Fellows	41
 Faculty Research	42
 Guiding Principles	45
 General Features	45
 Key Areas	46
 Biomedical Data Sciences	46
 Educational Analytics	46
 Cybersecurity	47
 Democracy	47
 Environment	47
 Business, Commerce and Finance	47
 Affiliations with other campus units envisioned?	47
 School of Data Science Satellites and Centers	47
 Centers	48
 Satellites	50
 Interactions with other Pan-University Initiatives	51
 Education - Curriculum	52
 Vision for degree offerings: names and descriptions	53
 Undergraduate certificate in data science	53
 Undergraduate minor in data science	53
 Undergraduate degree in data science	53
 Masters degree in data science (Grounds)	54
 Masters degree in data science (Online)	54
 PhD program	54
 Professional certificate in data science	55
 Research experiences for undergraduates	55
 Accreditation	55

IV. Faculty & Staff	55
Overview	55
What the DSI has in Place	55
Faculty and Staff Going Forward	57
Tenured Faculty	57
Tenure-Track Faculty	57
Tenure and Tenure-Track Appointments	57
Joint Appointments	58
General Faculty	58
Staff	59
Future Numbers and Growth	59
Endowments	60
Organization	62
Use Cases	64
Diversity and Inclusion	65
V. Students	66
VI. Resources	67
Physical facilities	67
Administration	68
Relationship with existing UVA administrative staff responsibilities	68
Plan for student advisement	69
Financial analysis and budget projections	69
Budgetary support	69
Plan for sustainability of the school	70
VII. Implementation Plan	70
Timeline	70
VIII. Evaluation	71
Train a responsible interdisciplinary workforce	71
Undertake leading-edge interdisciplinary, open research	71
Maximize societal benefit	72
Conclusion	72
Appendix A The School of Data Science - An Open Scholarly Ecosystem	73
Proposed Charge	73

Motivation	73
Impact on Other UVA Schools	74
What are the specific issues an open School will address?	74
How will the School address these issues?	75
Conclusion	77
Appendix B Support Letters from Deans, Vice President for Research, Internal Collaborators (Departments, Institutes, and Units), and DSI Advisory Board Members	78
Appendix C - Other Letters of Support	96

Executive Summary

The University of Virginia, through the largest gift in the University's history, has the opportunity to play a national and international leadership role in data science training, research, and service by expanding the already successful Data Science Institute (DSI) to become a School of Data Science (SDS). When first presented to then President-elect Ryan, he pointed out that a gift alone does not make a school. Particular concerns were sustainability and the impact on other schools of the University. Throughout 2018 and early 2019, we have crafted a proposal for the SDS that is financially and academically sustainable and that works in concert with all schools to enrich every student's experience at a time when our society is increasingly data driven.

The formation of the SDS also happens at a time of significant disruption in higher education brought about by, among other things, changing demands in the workforce, escalating costs in the face of diminishing public funding, and rapidly changing technologies. The SDS is an opportunity to respect the traditions of the past while at the same time providing a testbed for new modes of research, education, and service that challenge the status quo. That balance is reflected in what follows.

This submission to the Faculty Senate conforms to the Evaluation of New Schools template set forth in 2008 by the Faculty Senate Academic Affairs Committee. Sections I and II (Phase I) provide the context and justify the need for the SDS. Sections III through VIII (Phase II) provide a blueprint for how the SDS will be established, with special emphasis on faculty as leading stakeholders in the enterprise. Additionally, this document (inserted between Phase I and Phase II) addresses particular questions and concerns as defined and requested by the Faculty Senate Academic Affairs Committee (AAC) after discussion of the Phase I document.

This document does NOT include any new educational programs for consideration, those will be submitted separately as the SDS develops. This document provides a framework and vision for a new and innovative School of Data Science (SDS).

Background

The University of Virginia (UVA) has the opportunity to transition the Data Science Institute (DSI) into a School of Data Science (SDS). The process is facilitated through a gift of \$120M made to the University by the Quantitative Foundation. This proposal differs from recent past submissions for new schools in that the basis of the foundational educational program (i.e., the Master of Science in Data Science), research initiatives, and service have already been established at the University.

Planning for this opportunity began in early 2018. Several documents have been prepared, including a five-year pro forma budget, to illustrate the sustainability of the SDS. These documents form the basis of what is being submitted to the Faculty Senate for consideration based on the process outlined in the 2008 Academic Affairs Committee annual report.

To date, the following steps have been taken:

- Members of the DSI have been notified of the possibility of converting the institute to a school. They fully support this change.
- Discussions have been had between the past President, the current President, the Deans, and the Director of the DSI. All Deans wrote letters of support (attached in Appendix B) calling for the establishment of the SDS.
- Discussions have been had with faculty and chairs in departments potentially affected by the creation of the SDS, notably Computer Science, Statistics, and Engineering Systems and Environment¹, the three core departments teaching in the existing Master of Science in Data Science (MSDS) program offered through the DSI.
- On December 4, 2018, the Chair of the University's Faculty Senate was briefed on the upcoming presentation to the Board of Visitors (BOV) regarding potential establishment of the SDS.
- On December 7, 2018, the BOV was briefed by the President at which time they fully endorsed the concept of establishing the SDS.
- On December 10, 2018, the Executive Council of the Faculty Senate was briefed on the opportunity by the President, and the President discussed the BOV endorsement with all Deans. All Deans remained fully committed to establishing the SDS.
- On January 18, 2019, the President publicly proposed the SDS and announced a gift of \$120M to support it.
- The needs assessment (Phase I) proposal was sent to the chair of the senate Academic Affair Committee on January 25, 2019.

¹ The merger of Civil/Environmental Engineering and Systems and Information Engineering.

- The Phase I proposal was reviewed and approved by the Academic Affairs Committee on February 7, 2019. A number of questions were raised, which are addressed in this version of the combined Phase I (revised) and Phase II document.
- Subject to the questions being addressed, the Executive Council of the Faculty Senate, which also met on February 7, 2019, were unanimous in their approval that a School of Data Science was indeed justified.
- With acknowledgment of the issues addressed and documented from the AAC meeting and the recorded concerns from affected departments to be addressed in the Phase II submission, the Phase I document as submitted was approved by the full Faculty Senate on February 18, 2019.
- This document was submitted to the AAC for consideration on February 27, 2019.
- Comments from the AAC were received on March 10, 2019.
- Having addressed the comments a revised version of this document was returned to the AAC on March 15, 2019.
- A public meeting of the AAC on March 21, 2019 brought forth significant support for the School of Data Science from several department chairs and interested parties as reflected in the minutes.
- The document was further modified based on AAC input and sent back to the AAC on March 29, 2019.

I. Mission and Goals

It is first important to define data science.

Definition of data science

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from data in various forms, both structured and unstructured. Data science sits at the intersection of computer science, systems science, statistics, mathematics and information science. Conducting data science transcends traditional disciplinary boundaries to discover new insights, often by combining disparate datasets that would not likely be brought together otherwise.

With that definition, it is easy to see why the SDS will be a different type of school, dependent upon a very close working relationship with other schools of the University. This has been the mandate of the DSI since the beginning.

With this definition of data science, the mission of the SDS can be specified as follows.

Mission

- In concert with other University schools, train a diverse workforce at the undergraduate, graduate, and postgraduate levels to be *responsible* practitioners and leaders in an increasingly data-driven society.
- In collaboration with other schools, undertake leading-edge interdisciplinary, open research that is not likely to be undertaken in a single school and thereby catalyze research discovery through shared data and analytical techniques.
- Serve the University, the local community, the Commonwealth, the nation and the world by applying data science to *seek the truth* and *maximize societal benefit*.

No other institutional data science initiative known to us places such heavy emphasis on the responsibility that our students, faculty, and staff (our team) bear to ensure ethical and responsible use of the skills we possess to benefit society. We do not just teach these principles; they are (in the case of the DSI) and will be embodied in every action the SDS will take. As will be described subsequently, the SDS will have a specific committee, Responsibility, Diversity, Accessibility, and Transparency. That committee will examine and advise all new educational and research programs even while scholars will have academic freedom in research, publishing and teaching. We already have 1.5 ethicists on a team of 19 and this will be expanded as our programs grow. Their efforts are embodied in a Center for Ethics and Justice (described subsequently) which - as far as we know - is the first such Center/Department of its kind. Other universities and schools only offer a few scattered courses on ethics. We will be the first to have ethics and justice as the foundation of everything we do.

Similarly, no other data science initiative places such emphasis on working with *all* disciplines. While most of the funding to date in data science has come from finance/business and STEM fields, our vision is to establish and maintain academic diversity by providing education and conducting research in the humanities, social and behavioral sciences, and other fields where new discoveries can be made.

This mission can be translated into the following list of goals which are fleshed out through the remainder of the document.

Goals

- Overall

- *Purpose* - Over time establish an eminent faculty and staff and produce alumni who are leaders in data science for societal benefit and from which national and international recognition will arise.
 - *Interdisciplinary* - Enable and jointly achieve outcomes that are broadly recognized and measurable as *not* being achievable by individual UVA schools alone.
 - *Responsible, Diverse, Accessible, and Open* - Establish these criteria as foundational in all SDS teaching, research, and scholarship.
 - *Physical Presence* - Create a recognized physical home with satellites in each school, that exemplifies the 21st century version of the academical village - an ecosystem of people, ideas and infrastructure (both public and private).
 - *Virtual Presence* - Establish a school without walls - having measurable positive impact across Grounds and beyond.
 - *Relevance* - Remain relevant as technologies and economies change.
 - *Sustainable* - Be able to sustain and grow the enterprise regardless of local and global circumstances.
 - *Attractor* - Prove our ability to attract faculty, researchers, students, and staff to UVA.
 - *Retain* - Keep those we attract in a competitive marketplace.
 - *Giving* - Grow and sustain relationships with alumni who continue to engage with UVA both intellectually and financially.
- **Education and Training**
 - Establish acclaimed online and residential curricula that emphasize ethical principles and practical training.
 - Integrate with and support courses with a data-science component in schools and departments across Grounds.
 - Advise and collaborate with schools on new data-science educational initiatives.
 - **Workforce**
 - Train alumni to become responsible community leaders.
 - Encourage alumni to continue to engage with the SDS.
 - Hire and retain UVA faculty who gain national recognition through mentoring and teaching awards.
 - Hire and retain UVA faculty who become nationally recognized for innovative research.
 - Hire and retain UVA administrative staff who receive institutional and national recognition for building and enabling these opportunities.
 - **Research**

- Undertake data-science research in specific domain areas where we have the capacity to become world leaders in those areas.
- Be known for contributions to interdisciplinary research initiatives that achieve societal benefit.
- Achieve highly accessed scholarship that includes, but goes beyond, the traditional literature to include all parts of the research lifecycle - data, algorithms, software, and supporting documentation.
- Host the products of research through a Commons - an (as appropriate) open platform for all aspects of the research lifecycle.
- Place a strong emphasis on translational research - turning research into products and services.
- **Service**
 - Establish responsible collaborations with the private sector that achieve trusted long-term relationships.
 - Work with university leaders in developing a strategy for data governance and defining the necessary infrastructure to maximize the value to the Institution of its own data.
 - Continue to support/co-sponsor and expand events on Grounds and beyond, notably northern Virginia, related to data science, e.g.,
 - Women in Data Science (annual event ~250 attendees)
 - Hackathons (periodic event ~20-30 attendees)
 - Governor's analytical summit (one-time event with Gov. McAuliffe and broad private sector engagement ~ 350 attendees)
 - Datapalooza (annual event ~600 attendees from across Grounds and beyond - major engagement event with the community)
 - Lunch and Learn seminar series (~monthly with 20-75 attendees with speakers on topics as broad as property eviction data analysis to predictive approaches to materials science)
 - Fireside Chats (1-2 per year with distinguished speakers in data science, e.g., Francis Collins, Director, National Institutes of Health; President Ryan; Nate Haskins, Chief Data Officer, S&P Global)
 - Distinguished lecture series (1-2 per semester, e.g., Oscar Wood, Founder and CEO of NNData)
 - Expand what the DSI does in local, national and international communities; for example, forming more partnerships with sister institutions worldwide to exchange students, best practices, etc.

- Establish a more formal and strategic governmental relations effort at local, state, federal and international levels in order to enable more openness, transparency, and evidence-based decision-making.

How to prioritize achieving these goals and to evaluate our success is part of a strategic planning process that is under way. This work is being done in concert with the UVA strategic planning process (the DSI Director is part of that committee), so both strategic planning efforts are in lockstep. Our strategic planning now also includes input from the Academic Senate, and the future SDS organizational structure seeks to continue that relationship as described subsequently.

A major part of the SDS strategy in a data driven world is to determine not what to do but what not to do. The DSI Board has suggested that we think of growth in terms of a supply chain. Rather than chasing the ever-increasing demand, we should determine what we can supply to meet the intellectual and workforce needs while maintaining the quality and academic integrity of the University of Virginia.

What is the intellectual purpose of the school?

Data science does not exist in isolation but applies the various sciences that touch data to all existing disciplines, both STEM and non-STEM. The intellectual purpose of the SDS is to enable other schools to maximize the value of data and data analytics. As such, the SDS is an intellectual clearinghouse for faculty expertise, tools, data, and other resources.

Research - Team members (faculty, students, and staff) undertake research across traditional disciplinary boundaries. Their expertise is acquired both from other disciplines and generated internally. Existing DSI research projects exemplify this intellectual role to be expanded with SDS formation. Consider an example. The DSI played a significant role in enabling UVA to secure a Clinical Translational Science Award (CTSA) which is one national hallmark of a successful academic medical center. Data Science was a centerpiece of the proposal, with the DSI playing a key role that enabled researchers in other schools while furthering DSI faculty research agendas; in short, a win-win. A variety of similar outcomes across a number of disciplines could also be cited. Moreover, the DSI demonstrated its usefulness as a recruitment incentive in 2016-18, through joint appointments for faculty in computer science, biomedical data sciences, cybersecurity, and business analytics. The SDS will allow a significant expansion of this role.

Education - The SDS will significantly expand a faculty who can teach core concepts in data science as well as apply data science to all domains. By working in collaboration with other

schools and their existing courses while also expanding courses under the DS mnemonic, the SDS will contribute to the workforce needs of the Commonwealth and beyond.

Service – Within UVA, the SDS has a role in designing the needed institutional infrastructure and data-governance policies required by a data-driven organization to become efficient and to outperform its competitors. This includes student projects designed to help organizations and units (both research and administrative) use more data science methods to enable data-informed decision making. The SDS should not be responsible for central support services, such as high-performance computing (HPC) infrastructure and data management.

What body of knowledge does the school encompass?

The SDS, building on the DSI, will expand a knowledge base of existing data science experts covering as many disciplines as practical. To date, we have experts in systems, biomedical data science, digital humanities, finance, and business analytics. The SDS will work collaboratively with schools and departments to build out verticals (specific domains, such as transportation, energy, finance) in areas of determined need. This would include working with existing faculty across schools and making faculty hires at all levels (both jointly and within the SDS) to expand expertise in these areas. These verticals will necessarily include experts in the key areas of data science (computer science, statistics, math, ethics, and related departments) with an interest in those domain areas as opportunities for collaboration and building expertise in those departments. Beyond people, through an Open Data Lab, the DSI is providing access to data and tools to undertake replicable research, initially for capstone projects. Under the auspices of the SDS, the ODL will be expanded to provide capabilities not available through the library or elsewhere on Grounds.

How does the proposed school further distinguish the University of Virginia?

Table 2 illustrates that while UVA was early in developing a data science presence by way of the DSI (2013), others (116 academic institutions surveyed) have caught up. Indeed, some have gone ahead. A School of Data Science (SDS) will place us again in the lead and will enable us, through the innovative approaches outlined herein, to be a leader in responding to a data-driven society.

What degrees are to be offered?

No new degree programs are sought at this time. When new degrees and certificates are sought, they will be developed by the school's faculty and submitted to the Faculty Senate for review under the standard protocol. Further details on degrees and certificates already in place

at the DSI or to be proposed as a part of the vision for future educational initiatives, as well as how these student enrollment numbers were derived, are outlined in Section III. The numbers are estimates required to make a budget projection.

Table 1: Degrees Offered or Under Consideration and Anticipated Student Enrollments per Year
 (footnote marks in the table reference the more detailed descriptions below)

Degree type	19-20	20-21	21-22	22-23	23-24
Undergraduate certificate in DS ¹		100	200	400	700
Undergraduate minor in DS ²			TBD	TBD	TBD
Undergraduate degree in DS ³			25	50	100
MSDS (on-Grounds) ⁴	50	75	100	100	100
MS Dual Degree DS/MBA ⁵	15	20	30	50	50
MS in DS (Online) ⁶	20	100	200	200	200
PhD ⁷		4	8	12	16
Professional Certificate in DS ⁸	30	60	90	120	120

1: The undergraduate certificate in data science is currently being prepared for submission to the Faculty Senate as a 12-credit specialization which includes an identified set of core competencies and one elective.

2: The undergraduate minor in data science will be proposed after the undergraduate certificate has been developed and is expected to comprise between 18-24 credits with a practical component.

3: The undergraduate degree in data science is anticipated as a 4-year degree with an intentionally small cohort (estimated at 100 graduating per year starting in a steady state). The intent is that these would be additional students and not impact the undergraduate quotas of other schools. When put forward to the Faculty Senate the SDS proposal for an undergraduate degree program will address how the SDS plans for new enrollment within the context of the larger academic community, including internal and external transfers, and core liberal-arts education provided by collaboration with other schools. This planning will be done in consultation with the Deans and the Provost, taking into consideration the University's overall enrollment management strategy and enrollment growth plan approved by the Board of Visitors.

4: The on-Grounds MSDS is an enrollment expansion of the existing approved 11-month DS1 program, currently a class of 49 students. Our intent is to grow to two cohorts of 50 each by 2021-22 to maintain the cohesive and personal experience that students have from the point of the admissions process to the job placement programs.

5: The MS dual degree program is an enrollment expansion of the existing approved program with Darden from the current 20 students (7 graduating; 13 first years) to a complete cohort of 50 students. This program is an example of a collaborative effort to develop a program that raises the interdisciplinary capabilities of all students in both organizations. The dual-degree students participate on project teams and classes as a cohort as well as with students in both individual programs.

6: The online MSDS degree program is where big growth is anticipated. This could easily be expanded beyond current projections of 200 concurrent students (by 2021-22), assuming faculty could be hired to teach the material. This program is already under development in collaboration with Noodle Partners (<https://www.noodlepartners.com>). Applications opened in December 2018², and the first courses will be offered in summer 2019. The Noodle platform is being developed to integrate with UVA's existing systems for admissions and student tracking. Since UVA owns the software and content developed, it may offer other schools an existing infrastructure and approach for their own online courses should they choose to follow that path. In short, SDS will be a testbed for other schools -- a role the SDS is keen to play.

7: PhD program. The Provost approved a request by the DSI to develop a collaborative PhD program for review by the Faculty Senate and later SCHEV. This is independent of the creation of SDS. The DSI expects to bring the PhD proposal before the Faculty Senate in the second quarter of 2019.

8: Professional certificate in data science. The DSI already offers boot camps (short one-week courses tailored to specific audiences, typically professionals seeking data-science experience), and it is expected that a professional certificate will grow out of those offerings, notably in northern Virginia. Offerings will be developed in collaboration with the School of Continuing and Professional Studies (SCPS), with SDS faculty hires and joint faculty hires from interested schools to provide courses. The specific certificates to be offered, the duration of the programs, and number of students are yet to be determined.

What will be the nature of the research contributions of the members of the school?

Since data science potentially cuts across all disciplines, we will seek to establish research contributions in collaboration with strengths, or identified needs, in other schools. As such, many SDS faculty will have dual appointments (outlined in Section IV). Research areas already planned or actively pursued by the DSI (with parenthetical examples) include:

² <https://onlinedatasciencemasters.virginia.edu/>

- Open hardware (e.g., machine learning support for the artificial pancreas project);
- Financial modeling / business analytics (e.g., fraud detection in Capital One credit card data);
- Neuroscience informatics with an emphasis on neurodegenerative disorders, autism, and imaging (e.g., a portal supporting Alzheimer's researchers);
- Biomedical data sciences more generally (e.g., how immune cells in the brain can be used to address emerging questions about cell type and function);
- Cybersecurity (e.g., detecting and minimizing network intrusions on the UVA network);
- Educational analytics (e.g., correlations between student performance and mental health);
- Translational clinical research (e.g., reducing repeat visits to the emergency department);
- Ethics and justice (e.g., what tweets tell us about how urban populations interact); and
- Smart cities (e.g., predictive modelling of motor-vehicle accidents).

Research areas yet to be pursued, but for which we have identified willing collaborators, include:

- Environmental science (e.g., reducing energy consumption on Grounds);
- Digital humanities (e.g. making information from UVA library archives easily discoverable);
- Materials science (e.g., machine learning applied to producing better plastics);
- Global health (e.g., pandemic modeling and defining a suitable response, data sharing in public health emergencies);
- Normativity, otherwise known as the evolution of rule and law making (e.g., natural language processing of all Supreme Court opinions); and
- Quantitative analysis of democracy (e.g., analysis of social media from newly formed democracies to determine common trends and differentiators).

A more detailed review of collaborative research directions is described in Section III. This includes current funded and pending initiatives, specifics of existing research, and proposed research directions.

What services will be delivered to the Commonwealth of Virginia and the broader world?

Workforce - The DSI, through its MSDS program, now in its 5th year, has delivered a highly trained and competitive workforce with 100% employment rates in a variety of sectors (Fig 2). The proposed expansion through the SDS will expand the number of students (Table I) and the

number of degrees with a focus on broad interdisciplinary training. The programs will also address the various levels of skills gaps in the general workforce from analytics-enabled programs to data-science positions (Fig 1). The faculty hired to teach this content and conduct research will add to the Commonwealth data-science ecosystem.

Research Output - The majority of research thus far has been through the DSI faculty and their students, MSDS student capstone projects, and Presidential Fellow graduate-student projects supported through the DSI, the Office of the President, and the Office of Graduate and Postdoctoral Affairs. This work will increase in depth and scope with an expanded faculty and student base who have a strong interdisciplinary research agenda. We will also add an endowed program of visiting fellows (postdoctoral and faculty scholars) from UVA and elsewhere who wish to develop skills in data science. This will allow faculty from across the University to participate in supported development of data-science skills and research in their own areas of expertise and will facilitate collaboration between researchers. This is supported by the gift as an endowment to assure that this influx in new talent and collaboration among faculty internally remains a priority. By these means, the SDS will develop a diverse, high-quality research ecosystem.

Such an ecosystem has already led to a number of trusted relationships with the private sector involving companies with a significant presence in the Commonwealth. Capital One, Inova, S&P Global and MITRE are examples. An expanded research program is expected to attract human talent as well as companies to the Commonwealth.

Service & Community - The DSI has a good working relationship with the Charlottesville community through support of the Charlottesville Data Portal, TomTom Festival, Hack Cville and other community-services projects. DSI team members sit on various city committees, and the SDS will expand on those activities. This includes an increased emphasis on working with city data. A noteworthy example already in place is the examination of social media and traffic pattern data from the events of August 2017, with the goal of being able to predict and prevent such events in the future³.

The DSI has worked diligently to assist the Commonwealth of Virginia to develop data-driven practices and share data across agencies and with researchers in order to increase efficiencies and improve the health and safety of Virginians. This has included numerous projects working with state data through individual state agencies to address problems with a data-informed approach. An example is the development of tools to improve the efficiency and effectiveness of newborn screening⁴ in collaboration with the Division of Consolidated Laboratory Services

³ <https://dsi.virginia.edu/projects/analyzing-social-media-assess-extremist-group-threats>

⁴ <https://datascience.virginia.edu/projects/keeping-infants-healthy-across-commonwealth>

and the Virginia Department of Health. DSI team members serve on boards and committees within the Commonwealth to address these issues with an emphasis on facilitating data sharing and working with state data.

At a national level, DSI team members are engaged in a variety of activities that strengthen our presence in an emergent field. This includes serving on various editorial boards, working committees and the like, where participation is governed by the mission of responsible action to improve the human condition. These activities will expand with SDS formation and an increased number of team members.

At an international level, a sister-institution program already underway with Dar es Salaam University in Tanzania, the Charité Universitätsmedizin in Berlin, the University of Manchester, and the University of Zurich will be expanded. These affiliations offer exchange of students, faculty, and best practices.

II. Need for a School

Why a School and not an Institute?

As the DSI has expanded its activities, it has become apparent that there are strategic and procedural reasons why a School of Data Science makes sense. When the Faculty Senate voted to have a Masters in Data Science (MSDS) degree to be awarded by the DSI, they went part way towards the notion of a school. Prior to this authorization, only UVA schools awarded degrees. Subsequently, the Provost approved the hire of general faculty by the DSI - a further step towards school status. Finally, the Provost approved the attendance of the DSI Director at all Deans' meetings and annual Deans' retreats. In short, DSI is operating very much like a school, but the administrative structure and stature of an institute hampers its ability to grow, operate, and excel in an area where the explosive growth in demand for data-science skills is arguably unparalleled by any discipline in the University's history.

A school is needed to hire tenured faculty

It is hard to sustain a single degree program without tenured faculty, let alone the programs to be put forward to the Faculty Senate after the SDS is established. A stable tenured faculty is needed to teach our programs, provide long-term consistency, and remain cost-effective. Borrowing faculty from other schools is more expensive than direct tenured or tenure-track hires or general faculty with multi-year appointments.

A similar argument applies to research. Identified research programs are long-term continuous commitments to an area of study. These require tenured faculty to form the backbone and to provide a research environment where tenure-track faculty can flourish and gain tenure.

In short, a core group of regular faculty (tenured and tenure-track) who are 100% in the school (many with courtesy appointments elsewhere) are critical in driving its academic mission. To hire tenured faculty requires that the DSI become a school. There is no provision in University policy to do so otherwise.

Recognition

Data science has become fundamental to the academic, professional, and everyday life experience. Establishing a school for its study will signal to the world that UVA is at the forefront of excellence in this space. A school is recognized by the quality of its faculty, staff (team in SDS parlance), and the research and quality of students it produces. Quality faculty will only be achieved with the right incentives. Tenure or the promise of tenure is the most important.

A school expresses a commitment to the importance of data science to the future of UVA and potentially allows us to outpace other institutions in a critically important area of development. As described subsequently, UVA was at the forefront in establishing an Institute in 2013, but has now lost its competitive advantage, as there are many such institutes nationwide. If UVA desires a leadership role, we need to take the next step and provide the stature of a school. To understand where leadership lies today, consider the University of California at Berkeley. Berkeley has gone a step further by creating a Division of Data Sciences and Information. The new division will wholly incorporate the current Data Sciences Division, the School of Information (I School), the Department of Statistics, the Department of Electrical Engineering and Computer Sciences (EECS), and the Berkeley Institute for Data Science (BIDS). It will be led by an Associate Provost and Dean. (UVA is organized in schools, not divisions, but the plans here are comparably ambitious.)

A school provides national and international stature in data science. Certainly an institute can achieve a national and international standing in research, but the term “school” as defined by the Academic Affairs Committee⁵ of the Faculty Senate includes instruction, research, service, and a primary faculty, all of which are essential to the continued development of data science at UVA.

⁵

<https://app.box.com/embed/s/1f84096jtjjizutmp3x01kv5ijihhvm9x/file/95417954874?showItemFeedActions=true&showParentPath=true>

Process

Given the explosive of data science, an institute does not have the organizational structure nor physical infrastructure to support the needed research, education, and service activities.

Sustainability

Within a hierarchical academic structure where funding flows directly to the schools, it is impossible to sustain a large enterprise without equivalent stature. Witness our pan-University institutes, which are not sustainable without significant funding from the administration. A major reason being that grants which are made possible by virtue of the interdisciplinary nature of the institute do not see the F&A (indirects) flow to that institute. Rather those indirects flow to a school, which is an unsustainable situation.

Financial sustainability requires a diverse portfolio of revenue -- tuition, philanthropy, research, and private-sector engagement. All are enhanced by having a school with a stable tenured faculty and appropriate administrative infrastructure.

Demand for the academic program/needs assessment/market analysis

There is a vast array of literature on the demand for data scientists. Two brief examples:

- Forbes May 2017⁶ - [IBM predicts demand for data scientists will soar 28% by 2020](https://www.forbes.com/sites/louiscolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/#23e6144b7e3b)
- Bloomberg May 2018⁷ - [This is America's hottest job](https://www.bloomberg.com/news/articles/2018-05-18/-sexiest-job-ignites-talent-wars-as-demand-for-data-geeks-soars)

Beyond popular expressions of a seemingly endless need for data scientists is the more serious analysis performed by the Business and Higher Education Forum (BHEF), *Investing in America's data science and analytics talent - the case for Action*⁸. Figure 1 illustrates the demand for business people with analytical skills as well as data scientists across a broad spectrum of industries. These data were derived from over two million job postings.

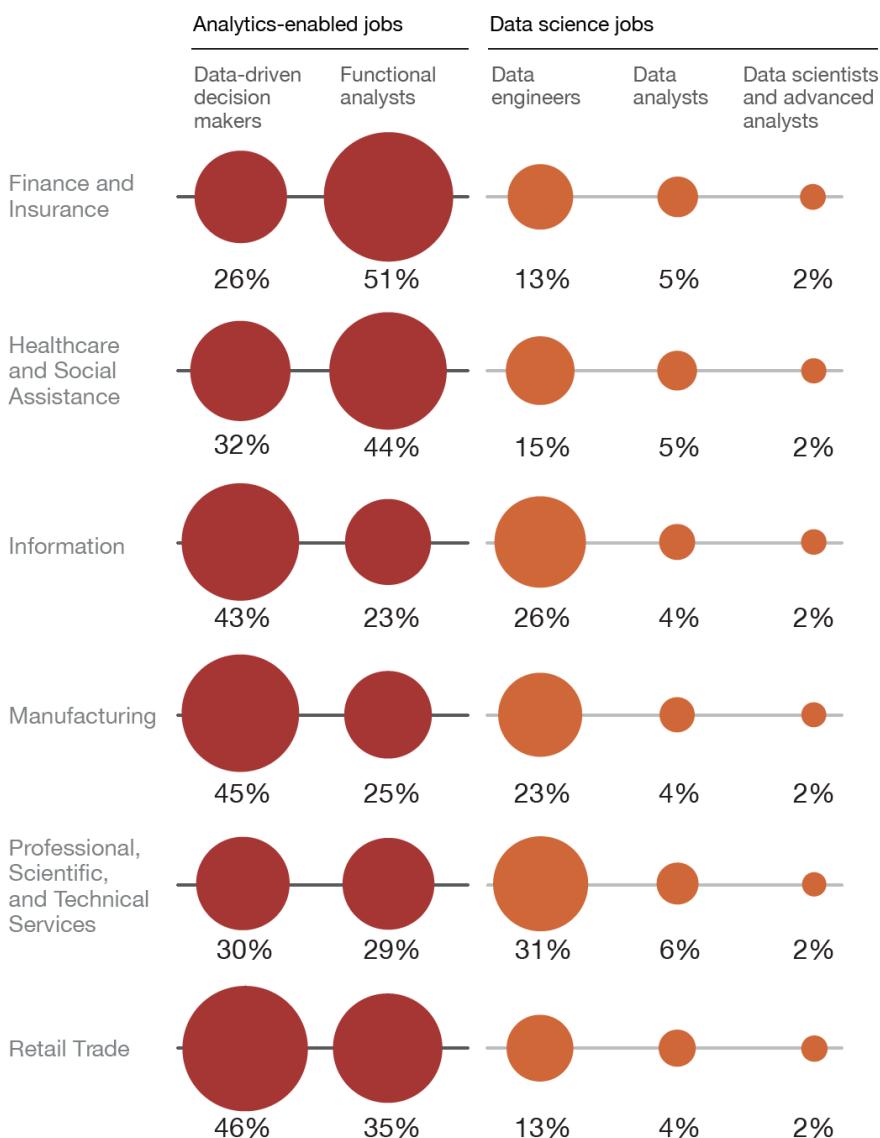
⁶ <https://www.forbes.com/sites/louiscolumbus/2017/05/13/ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020/#23e6144b7e3b>

⁷ <https://www.bloomberg.com/news/articles/2018-05-18/-sexiest-job-ignites-talent-wars-as-demand-for-data-geeks-soars>

⁸ <http://www.bhef.com/publications/investing-americas-data-science-and-analytics-talent>

Figure 1

Of 2.35 million job postings in the US.



Notes: Job category of analytics managers not shown. Totals may not equal 100%.

Source: PwC analysis based on Burning Glass Technologies data, January 2017.

Number of postings: Finance and Insurance (535,683); Healthcare and Social Assistance (100,000); Information (690,833); Manufacturing (237,484); Professional, Scientific, and Technical Services (511,947); Retail Trade (101,711).

Does the proposed school address a core need that UVA should meet to compare favorably with comparable universities?

To answer this question, first consider the offerings of other institutions, nationally and internationally. China is investing heavily in university data-science centers, and some countries, such as Switzerland, are creating national centers for data science⁹. Nationally almost all academic institutions are investing in data science, some heavily. Consider for example MIT's announced \$1bn investment in AI. We have picked three institutions to consider more carefully because they most closely match the scope of our ambitions. Berkeley seems the closest and indeed goes beyond a school to create a division of data science, which essentially houses multiple schools.

Berkeley Division of Data Sciences

- <https://data.berkeley.edu>
- Home: An interdisciplinary division at the center of an inclusive and adaptive organizational structure that will enable faculty and students to work together across boundaries.
- Integrates School of Information, Computer Science (EE), and Statistics, and existing Berkeley Institute for Data Science (BIDS).
- Degrees: Masters in Data Science (MIDS) program.
- Physical location: Includes a new teaching and discovery environment -- the **Data Science Commons**.
- More info: FAQ¹⁰, Org Chart¹¹

Columbia University Data Science Institute

- <http://datascience.columbia.edu>
- Home: School of Engineering and Applied Science (SEAS).
- Integrates by means of Centers: Foundations of Data Science; Cybersecurity; Data, Media and Society; Financial and Business Analytics; Health Analytics; Sense, Collect and Move Data; Smart Cities. Also: Columbia Data Science Society (CDSS).
- Degrees: Master in Interdisciplinary Data Science (MIDS), Certification of Professional Achievement in Data Science.
- Physical location: Engineering school with distributed centers.

⁹ <https://datascience.ch/>

¹⁰ <https://evcp.berkeley.edu/initiatives/division-data-science-announcement-frequently-asked-questions>

¹¹ https://evcp.berkeley.edu/sites/default/files/ddsси_орг_структура.pdf

Michigan Institute for Data Science

- <https://midas.umich.edu>
- Home: Based in the Advanced Research Computing (ARC) Office. Interdisciplinary focal point that includes a core faculty of 40 data scientists from various departments (statistics, biostatistics and mathematics, computer science and engineering, information science, and other experts). Also includes 200 faculty affiliates from over 60 departments.
- Integrates five research hubs: transportation, learning analytics, health science, social science, and music. Also convenes working groups based on research themes and funding opportunities that cut across traditional disciplines.
- Degrees: Undergraduate Degree in Data Science, Graduate Data Science Certificate Program, Masters in Data Science.
- Physical location: Unclear.

With proposed undergraduate and PhD programs as well as expanded MS programs, a robust research effort, and an emphasis on responsible data science and contributions to the community, we anticipate being highly competitive to all three of these top institutions. We can use this competitive advantage to attract the best and brightest students, faculty, and staff and create collaborative public/private partnerships.

As far as we can tell (Berkeley aside), we would be the first School of Data Science. As such, the SDS will have the distinct advantage of operating as an independent entity, working as an equal partner with all disciplines. This increases the interdisciplinary nature of what can be achieved since we are not emphasizing the work of a particular department or school, but rather emphasizing the work of a whole institution.

Employment opportunities for the school's graduates

Figure 2 illustrates the employment statistics from the 2016-17 DSI MSDS class. In 2018, the DSI teamed with SEAS to provide shared counseling across engineering (notably computer science) and data science domains. This kind of shared resource provides students across various schools with expertise in this area without a siloed creation of duplicative operational infrastructure and could be expanded to other schools upon SDS formation.

2017 Employment Statistics

Master of Science in Data Science

UNIVERSITY OF VIRGINIA
DATA SCIENCE INSTITUTE

The MSDS degree at the UVA Data Science Institute distinguishes itself with a focus on collaboration, project-based learning applied to real-world problems, and industry, academic and government partnerships.

To learn more, visit dsi.virginia.edu.

JOB TITLES

Machine Learning Engineer
Data Analyst
Data Scientist
Director of Data Science
Senior Statistician
Systems Engineer
Modeling & Simulation Specialist
Associate Consultant
Deployment Strategist
Sr. Associate Data Scientist

EMPLOYMENT OFFERS



CAREER TREK : 8 Students

San Francisco Bay Area

BIRMINGHAM DATA SITE

IBM

workday

NOTCH

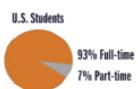
CCRI

HUMANGEOTech

2017 COHORT

• 47 STUDENTS • 66% MEN • 34% WOMEN • 83% U.S. • 17% INTERNATIONAL

EMPLOYMENT STATUS



SALARY RANGE

\$60,000 min. \$145,000 max.
\$91,515 average salary

SIGNING BONUS

\$5,000 min. \$15,000 max.
\$11,774 average bonus



★ Top three locations where MSDS Class of 2017 alumni are working.

● Locations where MSDS Class of 2017 alumni are working.

*Information based on 91% of 2017 cohort reporting data.

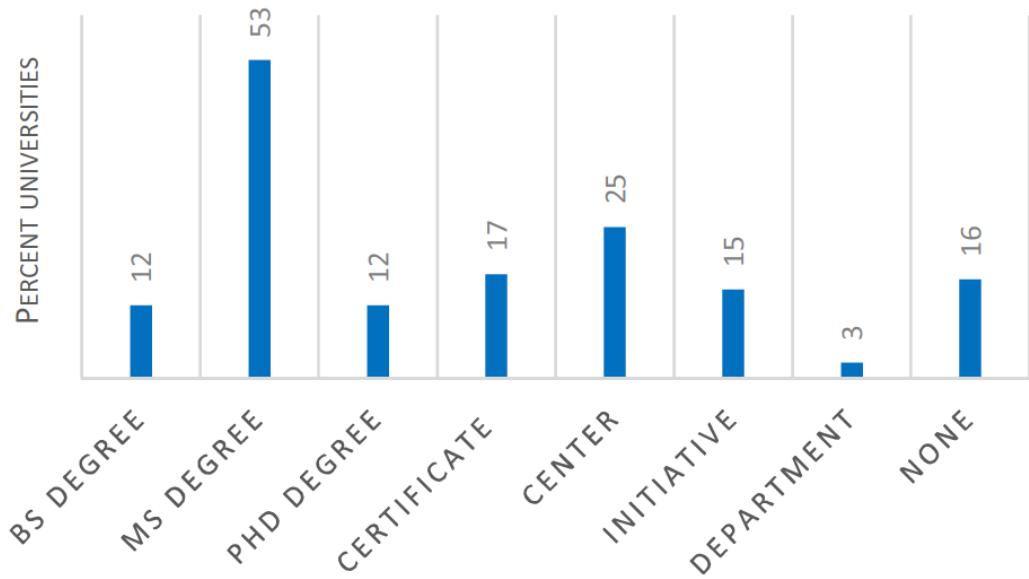
Figure 2 2017 DSI Employment Statistics

In summary, DSI graduates are fully employed in a variety of positions and geographic locations (with a concentration in northern Virginia) and over many employment sectors. This trend is expected to continue into the foreseeable future, even as the number of graduates substantially increases.

Review of competitive programs at comparable institutions

In 2011, the now-DSI Director was part of an initial advisory group convened by the Gordon and Betty Moore and Alfred P. Sloan Foundations to propose national developments in data science. In 2013, this led to the funding of three centers. Since then, a recent report released by the Sloan Foundation identifies another 113 data science initiatives in the US. The degree offerings from these institutions are outlined in Figure 3.

It is to UVA's credit that it began early in the evolution of data science within academia. However, it is clear from Fig. 3 that our programs need to expand if we are to retain a leadership role.



Source: *Moore Foundation, 2017.*

Figure 3 Data Science at Other Universities

Table 2 supports this statement through comparing the degree of faculty and scientific researcher engagement at the respective institutions. The expansion proposed by the SDS will propel us into a leadership position.

N	Center	Year launched	Space	Based in single departm/ college	Non-faculty managing director	Faculty lines	Data scientists	Postdocs
1	BU DSI	2012/14	✓		✓	✓	✓	✓
2	Caltech CD ³ JPL CDST	2015	✓				✓	
3	Columbia DSI	2012	✓			✓	✓	✓
4	Duke iiD	2013	✓				✓	✓
5	Harvard DSI	2017	✓					✓
6	JHU IDIES	2012	✓			✓	✓	
7	MIT IDSS	2015	✓		✓	✓	✓	✓
8	MSU CMSE	2015	✓	✓	✓	✓		
9	NYU CDS*	2013	✓			✓	✓	✓
10	NW DSI	2015			✓	✓		✓
11	OSU TDAI	2015	✓		✓	✓		
12	Stanford DSI	2014			✓			
13	UC Berkeley BIDS*	2013	✓		✓		✓	✓
14	UChicago CDAC	2000/18	✓					
15	UMass CDS	2015		✓		✓	✓	✓
16	UMichigan MIDAS	2015	✓		✓			✓
17	UNCC DSI	2012	✓			✓	✓	
18	URochester GIDS	2013	✓			✓	✓	
19	UVA DSI	2013	✓			✓		
20	UW eScience*	2008	✓		✓	✓	✓	✓

*Moore-Sloan Data Science Environments.

Table 2 Data Science Initiatives - Establishment and Offerings

Results from preliminary vetting of the proposed school

As stated above, the proposed school has been vetted with the BOV, University leadership (President, Provost, Provost-elect, EVP of the Health System, COO, VP for Research), the Deans, and the DSI Advisory Board. All are unanimous in support of creating the SDS as quickly as possible, while at the same time performing due diligence with stakeholders: faculty, staff, students, the private sector, and the community. Waiting too long to establish the infrastructure for this kind of initiative threatens a loss of competitive advantage.

SCHEV has provided preliminary information indicating what is required to seek approval. The SCHEV process is greatly simplified given that no new educational programs are proposed at this time and that SDS will operate via centers and satellites, not academic departments (outlined below).

Student target market

The target market for Masters students remains what it has been for the MSDS program and is provided in Figures 1 and 2. Beyond that, the proposed PhD program will train the next generation of data-science academics and industrial researchers. Finally, through collaboration with other schools, the undergraduate certificate, minor, and degree programs will train students with highly sought after analytical skills. (See columns 1 and 2 in Fig. 1.)

Impact on existing academic programs at the University of Virginia

The DSI Director has met with all the deans, a number of department chairs, and faculties to review impact on existing programs. They are satisfied that programs can be synergistic rather than competitive. The opportunity to work with SDS faculty and students would be an incentive for prospective students interested in closely related programs to come to UVA. Every care will be taken to maximize coordination. An example is business analytics. Darden and McIntire offer a joint MS degree in business analytics, while the DSI offers an MSDS and Darden and DSI jointly offer an MSDS/MBA. The deans of these schools and the DSI Director meet periodically to ensure the values and differentiating features of each program are clear to all stakeholders, especially to the students applying to these programs. The opportunities brought forth by both programs can strengthen each other and UVA's presence in Charlottesville, the DC area, and online. Sensitive to the fact that these differences may not be as easily apparent to outsiders, the teams need to work continually to define these differences and maintain close collaboration.

AAC Requests for the SDS Phase II Document

Responses from the DSI are shown here in red italics and incorporated into the Phase II document which follows these responses in Sections III onwards.

Responses are primarily related to what we are asking the senate to endorse at this time which is the framework of the School – organization, governance, etc. and not any new educational programs.

The Academic Affairs Committee looks forward to the Phase 2 proposal for the School of Data Science. Based on our review of the Phase 1 proposal, the committee would like to see a number of additional items included in addition to the standard elements of a Phase 2 proposal¹². Those additional items include the following.

Clarify the role of the Math and Statistics Departments: The Phase 2 documents should explain more clearly the role of the Math and Statistics departments in the programs of the proposed SDS. The intent here is for the phase 2 document to detail its vision as to how those two departments will interact with, and coexist with, the SDS.

¹² <https://virginia.app.box.com/s/n9ft1hj7ie5kigr0mvxe2163dml2b7km>

The intent is to maximize interactions between the SDS and Statistics and Mathematics Departments - as well as others - while recognizing that all have their own core missions. Stated simply, statistics and mathematics are applied throughout data science, but have fundamentally different research missions. Having said that, the SDS wishes to maximize shared governance and interdisciplinary activities with these departments. We are already doing so with computer science. To this end the SDS will:

- *Continue to work with the Dean of A&S to maximize the collaborative synergy that is already in place (see Appendix B).*
- *Reach out to chairs of Statistics and Mathematics on a quarterly basis to review opportunities and concerns.*
- *Invite members of these departments as representatives of the various oversight committees being established as part of the SDS. At this time, they are defined as (see Fig. 6 and Section IV):*
 - *Academic Executive Board which oversees all academic operations of the SDS.*
 - *Academic Affairs which oversees all on-grounds, online, and professional education programs. This expands what is already in place for the DSI MSDS program.*
 - *Strategic Affairs which oversees all strategic initiatives of the SDS, including outreach and partnerships.*
 - *Research Affairs which oversees all research initiatives of the SDS.*
- *Offer the opportunity for joint appointments at all levels between SDS faculty and Statistics and Mathematics at both term and tenure levels.*
- *Have the Academic Executive Board consider joint hires proposed by Statistics and Mathematics.*

The greatest concern from these departments seemed to be impact on student numbers taking degrees in statistics and mathematics. Since no new degrees are proposed at this time, but rather a governance structure from which new degrees will be proposed, being part of that structure will hopefully alleviate concerns when new degrees are proposed. We would like to emphasize the potential opportunities the SDS will afford through expanding research programs, attracting faculty talent, and expanding collaborations with other schools and sectors outside of the university. Involvement in the SDS governance structure will facilitate these possibilities for the Departments of Mathematics and Statistics.

Addressing other departments' concerns: A concern is how the SDS will affect the other departments, especially Math, Statistics, and Computer Science. Those three departments will each provide an (approximately) 100-word paragraph summarizing their concerns (they are included at the end of this document). The Phase 2 proposal should directly address these at some point. In a separate document (which can be an email), the locations where these are addressed should be quickly summarized so that they can be found easily. The intent here is to directly address any concerns that these three departments have, which may (or may not) be similar in content to the previous point.

This document is that "separate document," and we will include it as part of Phase II to create a complete record. We suggest that when this process is concluded that the Senate revisit the

process for new School consideration in the light of this first use of the process. We are willing to help in this regard.

Mission statement: The mission statement should be updated to match the comments made in the Phase 1 AAC meeting. The notes from the February 7th AAC meeting should be used to guide this update. The essence is that in addition to seeking the common good we need to be sure to seek the *truth*. The SDS mission also makes clear the School's commitment to the public good through the cultivation of "responsible practitioners". The committee would welcome greater clarity, and specificity in the Mission section as well as throughout the Phase II document, around how the school intends to enact its stated mission of placing "emphasis on the responsibility that our students, faculty and staff... bear to ensure ethical/responsible use of the skills we possess to benefit society."

The distinction made between "truth" and "public good" is addressed in our revised mission statement and throughout the Phase II document. With regard to "ethical/responsible" we think this is where the SDS can distinguish itself based on what the DSI has accomplished already and that this is an important part of the philosophy of the University. More specifically what is outlined throughout the Phase II document is:

- *The DSI has 1.5 ethicists on our faculty, and we have established a Center for Data Ethics and Justice that will be expanded under the auspices of the SDS. The appropriate application of data science will come under closer scrutiny in the next few years. We will undertake research and training to address issues that rise.*
- *Our MSDS program has a core course in ethics. Such course offerings will be part of any new programs put forward.*
- *A track record of data ethical/responsible use of data will be further emphasized in our updated faculty handbook and in hiring.*
- *We look for the same in students seeking admission to our MSDS program.*
- *Ethical/responsible use of data is embodied not only in our teaching of specific courses, but in all we do. For example, a responsible part of capstone and presidential fellows' projects is a consideration of the implications of data use from an ethical/responsible perspective. This includes listening to the community and having an ongoing dialog with the community of the role of data science in society. This is already happening through public lectures, the DSI's participation in TomTom, and more. The SDS will expand this engagement.*
- *Since its origin, the DSI has sought to establish and participate in an ongoing, interdisciplinary conversation on the nature of truth and knowledge in the context of data science. Through the Center for the Study of Data and Knowledge, one of the DSI's original Centers of Excellence, and through collaboration with the Humanities Informatics Collective, the center has been actively pursuing questions ranging from algorithmic bias to the nature of science in the era of neural networks. As a school, we anticipate expanding our curriculum by collaborating with humanities departments to develop courses in critical data studies and other areas that apply the lenses of history, culture, and philosophy to achieve a fuller understanding of data as a human phenomenon.*

Faculty and tenure: There needs to be a guarantee in the Phase 2 documents that the majority of full-time faculty hires will be hired into tenured or tenure-eligible positions. The academic freedom that tenure protects is essential to the proposed School's ability to fulfill its mission for ethical scholarship in service to the common good. The committee realizes that there may be some non-tenured positions, such as the Academic General Faculty at UVa¹³, but they still have a form of job security, and their existence does not preclude tenure-track faculty. This should be included in an entire section on faculty. The AAC committee considers the possibility of no tenure in the School to be a complete deal-breaker all by itself. Thus, we request letters from the President and/or Provost-elect affirming that tenure will be offered to the majority of (full-time) faculty.

It is our intention to appoint tenured and tenure-track faculty. We will not be able to hire the best scholars nor sustain the SDS without making tenured appointments. The exact distribution of faculty without term vs those with term (general faculty) is the Provost's decision. We have discussed tenured appointments and their ratio to general faculty with the Provost-elect and include a letter from the President and Provost outlining their response.

We recognize and respect the importance of academic freedom afforded by tenure. We also recognize that different disciplines consider tenure differently, for example in relation to clinical faculty in the medical school. Under the auspices of the DSI we discussed with the Provost over a year ago and have approval for the provision of general faculty (teaching, research, and some combination) with 12-month appointments as well as 9-month appointments. As such, general faculty have already been hired into the DSI. The DSI already has a faculty handbook approved by the Provost's office for the hiring of general faculty and that will be expanded to cover tenure and tenure-track appointments. Details can be found in section IV below on Faculty & Staff.

Faculty and tenure redux: A vexing issue at UVa is how to support cross department/school tenure. For the type of school being proposed it would be ideal for this sort of tenure process to occur -- i.e. each department/school jointly put up the faculty for tenure. To date this sort of thing has not been particularly successful at P&T. The current paradigm is fairly insistent that a department owns a particular faculty member; this is likely driven mostly by financial elements. With the amount of pan-University collaboration being envisioned, how will the school handle this issue?

The notion of a primary and secondary department/school precludes true interdisciplinary research and education. The SDS is an opportunity for the University to try a new model. The faculty member always feels more beholden to the primary department where his or her major appointment lies. Moreover, the UVA administrative structure perpetuates this, with indirects flowing to primary departments. We propose a number of 50:50 appointments where the appointment is truly shared. That is, a special P&T committee will be assembled from the two departments/schools. Making such appointments will be reviewed with the respective deans/department chairs and approved well in advance of making offers. This might have the desirable side effect of bringing two departments/schools closer together. It is noted that a reverse model already occurs. Biomedical Engineering is a department that sits at the interface of two

¹³ <https://engineering.virginia.edu/sites/default/files/common/offices/deans-office/policies-bylaws/2017.7%20SEAS%20Policy%20on%20Academic%20General%20Faculty.pdf>

schools and it an undoubted success story. This is addressed in the letter from the President/Provost.

A New Kind of School: A core ambition of the School is to be a new kind of School, a highly collaborative “school without walls,” deeply connected to, and even embedded in, other schools across Grounds. This ambition is critical to the school’s identity and indeed its very success. Yet this model runs contrary to both long-standing University precedents and the financial incentives of the University Financial Model¹⁴. How will the President and future Provost manage the School and its relationships with other Schools to allow it to fulfill its uniquely collaborative mission? As part of this, how will overhead be shared between collaborating units? The committee would like to see a statement from the President and/or Provost-elect (perhaps included in the same letter that addresses faculty and tenure, mentioned above) that acknowledges this context, and that outlines a vision or management philosophy for how this new collaborative model for the School will be maintained in an enduring way in the face of these significant contrary pressures and precedents.

This letter is included. Speaking for the DSI, this sharing model is already in place. DSI has a tenure-track faculty member who is 50% paid by the DSI and 50% by Public Health Sciences (PHS). A memo of understanding between the department chair and the DSI Director insures overhead is shared appropriately. PHS does not see this as a loss of 50% overhead, but rather a gain of 50%, as this person would likely not have come to UVA if it were not for the joint appointment. Indeed, the ability to recruit top faculty is enhanced by the presence of the DSI, and we believe more so, given the level of interest in a School of Data Science.

Support of the deans: The Phase 1 proposal stated, “all deans wrote letters of support calling for the establishment of the SDS.” The committee requests that these be included with the Phase 2 document. If not appropriate for the body of the Phase 2 document, these can be attached as an addendum or appendix at the end of the Phase 2 document.

The letters are attached in Appendix B as are a number of other relevant support documents (Appendix C) which are referenced in the Phase II document.

Pan-university collaborations: The committee would like to see details that address the envisioned collaboration with a range of entities throughout the university. As *only one* example, we use the existing digital humanities programs and institutes at UVA, which are longstanding and very distinguished. For that example, the committee would like to see some specific examples of what these collaborations might potentially look like. Specific programmatic commitments are not expected, but rather a vision or set of possibilities. Letters of endorsement from the leaders of these organizations, such as (for this one example) Allison Booth (Scholars Lab), Worthy Martin (Institute for Advanced Technology in the Humanities) and John Unsworth (the Library) would be expected. That is but one example, and the committee would expect a range of such examples in the Phase 2 document.

Appropriate letters are included (Appendix C) and referenced in the text. Table 7 provides a number of collaborative activities across some schools and Tables 4 (Capstone Projects) Tables

¹⁴ <https://financialmodel.virginia.edu/>

5 (Presidential Fellows Research Projects) and 6 (Research Grants) further illustrate the degree of collaboration.

School, not degrees: The Phase 2 document should make it clear what is being approved in the Phase 2 document: the School itself, and not the degrees, which are going through a separate approvals process. That being said, information on the various degrees that the School expects to offer should remain in the document.

That we are seeking endorsement only of the formation of a School of Data Science has been made clear throughout. Part of that endorsement is that the existing DSI MSDS program would now fall under the auspices of the SDS. It has been made clear throughout that the approval of new programs is not sought. A description of anticipated degrees are included (Faculty Senate requirement in the template document provided), but they will be subject to a separate review process if and when they are sought.

Collaboration redux: How are the School's programs going to be coordinated with existing programs? (Examples include the new data analytics programs we approved last year, one for the School of Continuing Education, the other for Curry/Darden?) How are these new faculty to be compensated? Will they receive significantly accelerated salaries and benefits? Will this contribute to the already significant salary differences between the schools?

Details of coordination with existing programs are given in the Phase 2 document. The DSI Director/Acting Dean meets with the Deans and associates from Schools where there might be overlap with any program and reviews these issues from the point of view of the student experience, faculty resources, and financial implications. Of the examples cited, the MS business analytics program is between Darden and McIntire, while the DSI offers an MSDS, and DSI and Darden jointly offer a dual MSDS/MBA. An MOU was drawn up between the DSI Director and the Deans of Darden and McIntire which defines a roadmap for the respective programs and periodic review to minimize overlap. Similarly, the DSI and SCPS are working together on offering executive-education certificate programs in northern Virginia sharing resources and expertise.

Compensation is set by the Provost. Demand for faculty in data science is extremely competitive. Compensation is therefore likely to be favorable, but within our pro forma budget. This is why the SDS wishes to train the next generation of data scientists, some of whom will enter academia. This is no different than in other areas of high demand such as engineering and clinical practice.

Additional details: There were a number of items that were not described in the phase 1 proposal (as expected), but the committee would expect to see in the phase 2 proposal. They include:

- How SDS developed their enrollment projections for the various programs detailed (e.g. 1,000 certificates a year). This number seems very high from one point of view (a huge portion of the undergraduate student body) and very low if 59% of employers prefer it as was mentioned in the AAC phase 1 meeting. How were these estimates created?

These estimates (and they are only estimates) were based on what we observed at other institutions which already have these programs. Some numbers had to be estimated in order to make the budget projections required by the President before he would approve moving forward with the SDS. Final numbers will come through consultation with potentially affected schools and the Vice Provost for Academic Affairs. The organizational structure of the SDS insures that considerations of affected schools/departments are taken into account in developing any new program. The certificate in data science is a good litmus test and is being developed with representation of all schools in a committee co-chaired by the Vice Provost for Academic Affairs and the DSI Director.

- A discussion of where these majors/minors/certificate will be drawn from as minors and certificates are a direct draw from other programs.

Students will be drawn from across Grounds. Since the majority of students would be taught in courses already taught in different Schools (who will receive 75% of the revenues under current funding models) assuming they are willing, none will proceed without having those courses approving the program.

- Some analysis about what proportion of certificate receivers will become minors (in the proposal today the difference appeared to be only 4 credit hours).

The minor is estimated to be from 18-24 credit hours and that number is not known at this time and would be put forward as part of a program proposal. The only point here is to indicate that a minor is anticipated based on student demand and will be worked out with all the schools.

- A discussion of what certificates/minors will involve for majors that already require classes that would count towards a SDS degree. For example, would a stats major already qualify? How many of their courses count? Would every department have its quantitative methodology course/sequence count? How would this work for graduate students across the various Schools? What process would be put in place to create these rules? This discussion is germane here because it gets to the target market discussion and not simply to specific curricular concerns.

These questions are being addressed now for the undergraduate certificate in data science. The process will be applied to other programs. It process involves all schools (graduate schools are participating too) having representation on a committee overseen by the office of the Vice Provost for Academic Affairs and the DSI Director. Since data science touches all academic disciplines, shared governance is inherent in all programs to be put forward by the SDS. A major SDS contribution is an understanding of the target market which is so important to these discussions. The DSI/SDS has career counselors and is in constant touch with an expanding marketplace of career options for our students.

- An explicit discussion of the tradeoffs of a 4 year undergraduate vs 2 in the college & 2 in SDS. Again, this discussion is germane here because it gets to the target market discussion and not simply to specific curricular concerns.

This may change as we consider an undergraduate degree submission, but for now the intent is to have a 4-year degree in SDS with limited or no transfers. This enables us to control the numbers of students graduating and minimize any impact on other schools, notably the college, losing students. Having said that, if schools are willing, the majority of lower division classes will be taken in other schools, generating revenue for those schools. The data-science undergraduate degree will create UVA data science that speaks to the features of the SDS - practically and ethically trained.

SDS will also consider that undergraduates students may have a minor or major in other schools. This would encourage a cross-disciplinary well-rounded education and would ensure that SDS students don't silo themselves.

- A discussion of how the ethical aspects of the degree will be baked in structurally. Without these structures, ethics may not actually constitute a distinctive aspect of the school.

Our Center for Ethics and Justice will lead our undergraduate educational efforts, with oversight for the SDS at large coming from the Responsibility, Diversity, Accessibility, and Transparency Committee. Further faculty will be selected based on their varied contributions to the ethical underpinnings of data science. This foundation will insure that every course and every research initiative has thought through the ethical consequences of what is being undertaken.

Faculty governance structures: The proposed School is expected to have a number of dual appointments, joint appointments, courtesy appointments, etc. In addition, there are likely to be tenure-track faculty, non-tenure-track faculty (i.e., AGF), adjuncts, etc. While these concepts exist in all Schools to some extent, the collaborative nature of this School implies that they are likely to be a much more significant issue compared to existing Schools. A discussion of the governance structures to be put in place at the School to manage this complicated setup will help clarify issues surrounding governance. For example (one of many), how would joint hires work -- would they be in the SDS or the "other" department? This would likely fit into the suggested section on faculty (although the committee is neutral as to where in the document it appears).

This is discussed above and is discussed in the phase II document under faculty and staff.

Comments from three departments

As part of the second point in this document, the three departments judged to potentially be most affected by the proposed school -- Statistics, Mathematics, and Computer Science -- were asked to submit a short series of items to be addressed in the phase 2 proposal. These are included here.

Department of Statistics

1. Many (if not most) of SDS's predicted "100 majors/yr" are likely to come from existing majors (especially Statistics, which is a huge component of most people's view of "Data Science"). How will SDS avoid decimating existing undergraduate programs?

Limiting numbers of undergraduates will help, but the main objective is to develop programs collaboratively so that "decimation" does not occur. This begins with a realization of the role that data science will play in the future of the University. The phase II document has numerous examples of collaboration in education and research.

2. Collaboration is key to SDS success in both research and education. How will SDS ensure collaboration with faculty and students across all units (more successfully than DSI), and avoid duplicating research and teaching faculty & resources that exist already at UVA?

Collaboration is occurring successfully with other departments, as described in Tables 4 and 5. The shared-governance model described in Section IV helps define how collaborations will be expanded.

3. The "risk" of SDS becoming a "silo" was raised. What steps will SDS take to avoid that risk? Example: form a Strategic Steering Committee comprised of faculty from the entire University with expertise in areas related to data science.

The major functional modules in Fig 6, notably the SDS Academic Executive Board, Academic and Research Affairs invites membership from across the university. Such shared governance will reduce siloing as will true joint appointments as described in Section IV.

4. SDS Research mission needs to be better articulated. (Perhaps the proposed Steering Committee could help them.)

This is more extensively discussed in the Phase II document.

5. Incoming undergraduates should apply directly for admission into SDS as they do for CLAS & SEAS. Otherwise, CLAS & SEAS operate in "service" roles as "filters" to SDS.

This is the intent, but it is not a part of this submission as the undergraduate program submission will be considered under a separate review process.

Department of Mathematics

There are already departments in the university, e.g. statistics, engaged in data science work, and other departments, e.g. mathematics, which would be the natural home for research into other emerging areas of basic research (e.g. topological data analysis, tensor networks). Thus we would like to see structures in place that will encourage cooperation and discourage competition and "siloing" of the SDS. We recommend the formation of a steering committee, which would include faculty from departments in other schools that have an interest in collaboration/joint appointments with the SDS. The steering committee would be charged with recommending directions for future hires and forming search committees across schools as needed.

Such steering committees are built into the organization of the SDS as shown in Figure 6 and discussed in Section IV.

Department of Computer Science

Computer science would like to ensure that a School of Data Science is structured for success and in a way that benefits the entire university. This raises several important questions that we believe should be addressed in the charter:

We appreciate the collegiality and cooperation of CS.

- Appropriate balance of fundamental and applied content in both its research and educational programs

This is a key point that was raised when the DSI Director met with the CS faculty recently. It is a point that the DSI has been considering internally for some time. The proposed organizational structure (section IV) that includes representation by CS in educational and research programs will insure that as such programs move forward, the appropriate complementarity will be established to the benefit of all stakeholders. The Departments of CS, Statistics, and Mathematics provide much of the theoretical underpinning for data science, and all disciplines across Grounds provide the applications which is where much of the emphasis will lie. Stated another way, driving applications will foster close collaborations with Computer Science, Statistics, and Mathematics to achieve learning outcomes and a comprehensive research agenda.

- Appropriate balance of TTT and AGF

This is discussed in Section IV. As stated there, the appropriate balance is defined by the Provost. The desire of the SDS is to create a balance that attracts the highest quality faculty to achieve the SDS mission and provide sustainability.

- Critical mass of faculty and scholarship to be a top school internationally, as well as an effective nexus for collaboration within UVA

It should be apparent from all that is discussed in the Phase II document that a critical mass will be reached.

- Joint hires--financial aspects as well as assignment of responsibilities and accountability

This is underway and expands on joint hires made by the DSI. We are looking at a joint hire with Curry and another with the College. These will be true 50:50 hires and will require a shared P&T process which we are establishing with the Provost's office.

- Encouraging/enabling faculty from other schools to participate in satellites/centers; achieving critical mass to make these top groups in their topic areas.

If early indicators from the satellites in Curry, the College, and SOM are indicative, encouragement is not necessary. Moreover, the Deans of these Schools have been supportive in making this happen. See their letters of support in Appendix B.

- Governance: hiring priorities, creation/administration of satellites/centers, curricula, etc. If the SDS is to be a cross-cutting, synergistic school with strong partnership with other schools, it is important that stakeholders from other schools have a voice in all these decisions

Agreed. Defined in the organizational structure given in Section IV.

- Resourcing hiring to meet DS teaching needs—especially re startup

The DSI is hiring a personnel manager to handle hires and we have the budget to bring on new faculty to kickstart the teaching and other needs at startup.

III. Academic & Research Programs

We have broadened this category from the Phase II Faculty Senate template to reflect that the School of Data Science (SDS) will be a leading education, research and service school. We begin with further elaborating upon the founding principles that were presented and endorsed by the Deans and were introduced in the Phase I document (Parts I & II above).

Founding Principles

The University of Virginia (UVA) has a rare opportunity to create a school that embodies all the principles that has made the University great, while at the same time creating a forward-looking exemplar for what the future will look like in higher education, a digital future where traditional norms alone no longer limit the training our students need, nor the research that will better the world, nor what will sustain the organization. We propose several founding principles:

- **Be responsible, diverse, accessible and open.** This is, above all else, what defines the SDS. We believe that, in the years to come, how data science is used will rightly receive great scrutiny. The SDS is committed to producing professionals who are more than great scholars, teachers, and industry leaders, but who also embody principles of responsible leadership in how data are used and for what purposes. We will hire faculty who embody these characteristics and who will be evaluated on their commitment to these ideals. Those faculty will then train students to think carefully at all times on the ethical consequences of their actions. This will not be achieved simply by delivering courses in data ethics, but by encouraging debate on the ethical consequences of proposed actions. We will make constant consideration for responsible data science part of our daily work habits. Students will learn by consistent example. Tied closely to this approach is the desire to be diverse, accessible, and open. What is implied by the SDS being open is defined in detail in Appendix A. The most diverse workforce is the most productive because it brings varying perspectives to every issue. The DSI has already shown this in terms of gender and ethnic makeup as well as background, but we need to do better. Accessible and open applies to all stakeholders associated with the SDS - faculty, staff, students and the products and services produced by the SDS. Open is the default, where “open” implies, as much as is financially and legally viable, free and unencumbered use of all we produce by all stakeholders.
- **Be constantly strategic and nimble** given a fast-changing demand, while remaining aligned with the emergent UVA strategic plan. While this newly established field of data science will inevitably grow and flourish, what it will take to remain at the cutting edge will shift over time. The SDS needs to allow for this and be constantly reviewing our strategy and resources (human and otherwise) to meet these changes. This will be reflected in what and how we teach, the kind of research we undertake, and how and with whom we engage in the private sector. Whom we hire, the type of contracts we offer, and how we organize ourselves have to recognize this rapid change. Simultaneously, the purpose of a School of Data Science is to teach foundations of data-

science principles and not simply a suite of tools that can quickly become outdated. Therefore, as the SDS evolves, close attention will be paid to not following the hype of the movement, but instead developing strong foundations to enable adaptation to tools as they become available and more broadly used.

- **Be sustainable.** From a financial perspective, do not overreach. This is a hallmark of the DSI and reflected in the attached pro forma budget where expansion is tied to a diverse portfolio of available resources - endowments, gifts, indirects on grants, tuition revenue, contracts with the government and private sector, royalty, and licenses. Growth will be carefully managed within the constraints of available resources (financial and personnel). From a capabilities perspective, this means having a core tenured faculty, a satisfied staff, and a quality trained student body.
- **Be interdisciplinary**, benefiting all schools, pan-University institutes, Wise, and NoVA. This is the mantra of the DSI and woven throughout all that is written here for how the SDS will operate. Interdisciplinarity is part of the fabric of the emerging field of data science. Particular sensitivity must be given to departments which are closely aligned with data science, notably Statistics, Computer Science, and Mathematics. Data science exists at the intersection of these and other departments and will benefit greatly from teams with interest and expertise in these and other areas.
- **Be a school without walls.** Notable here are the satellites and centers, described subsequently, which are largely developed and run by faculty: the former with joint tenured appointments; the latter with SDS tenured appointments.
- **Be team-driven, not individually driven.** The DSI's success has come in part through operating as a team. There is minimal distinction between faculty and staff, and there is a focus on the development of leaders in all areas as the school grows and evolves. All have a say in major decisions. While it will be harder to sustain this in a larger organization, we are structuring ourselves to continue this tradition as it will enable us to bring in and keep talent in the school and the University. Our organization reflects this and is described subsequently.
- **Strive for quality, not quantity**, in education and research. We will never completely meet the ever-expanding demand for data science workforce, so we must determine what to supply. This fits within the academic mission of the University to strive for and focus on academic and intellectual rigor rather than following a moving target to churn out massive quantities of graduates in a particular area. Our strategy is to keep on-Grounds numbers of students modest within the SDS and to facilitate training in other schools to help with any expansion goals they have. Notable in this is our relationship with School of Engineering and Applied Sciences (SEAS) and the Computer Science Department in particular. In contrast, with the online Masters in Data Science (MSDS) and professional certificate programs, we expect a larger growth to address accessibility and flexibility in this important area of upskilling for working professionals.

- **Be innovative and translational** through new forms of engagement with the private sector, government, NGOs, local, state, and national and international partners. While being vigilant with respect to conflicts of interest and responsible data science, we plan a richer engagement with the private sector that builds upon trusted relationships already in place or yet to be developed. Those relationships were founded, in part, through our MSDS capstone projects across the private sector, and we plan on building those relationships, with the support of the Office of Advancement (through a joint hire), to include longer-term joint research projects. Those projects will in turn bring private-sector researchers into the SDS to work directly with our team and our students. This creates a direct interface for ensuring that what we teach remains relevant by collaboratively addressing real-world problems and exploring the types of employment opportunities available to our students. Beyond the private sector are relationships with government (local, state, and federal) and NGOs that can provide a lot of data, if properly shared and protected.
- **Serve** our institution, our community, our Commonwealth, and our nation. Service does not mean provide computer and data services typical of IT support. It means using the skills associated with data science to improve the wellbeing of our communities. So for example (as the DSI is doing now): proposing data governance and using the institution's own data to improve the wellbeing of our students and improve the likelihood of faculty receiving grants; working to improve the Charlottesville open data portal; serving on Virginia Information Technology Agency (VITA) committees; serving as editors for emergent data-science journals, etc. Service also means listening and engaging with the concerns of the various communities.

Fit with UVA's mission

How will the new school fit within the context of existing academic programs?

This is discussed in the context of research, education (notably the curricula), and service.

Research

Thus far, research has evolved with the growing faculty of the Data Science Institute (DSI) and comprises the capstone projects undertaken by the MSDS students, research undertaken by the Presidential Fellows, and research undertaken by faculty members and their students and postdoctoral fellows within the DSI. Each type of research is described briefly with emphasis on how that research will expand and change as part of the transition to SDS.

Capstone Projects

These are research projects undertaken by DSI MSDS and MSDS/MBA dual-degree students in teams of 2-4 for a period of 9 months during their 11-month MSDS. Projects are sponsored by the private sector, government, NGOs, and academic research labs (the client). Students are matched to projects using an algorithm that maximizes the connection between their interests and the available projects. Students are exposed to all aspects of the research lifecycle, culminating in a conference presentation and a scientific paper. There is also an emphasis on

telling the story with the data, explaining why the data and results matter, as well as how they were achieved. This makes the data-science lifecycle more accessible to the client and others interested in the project. Example projects reflecting the breadth of what is covered is provided in Table 4. More details of the projects undertaken can be found on the DSI webpages¹⁵.

Table 4 Example MSDS Capstone Projects

Discipline(s)	Economic Sector	Title
Digital humanities, STEM, social and behavioral sciences	Communications	Improving access to open science
Engineering, applied science	Transport	Preventing fatal motor vehicle accidents
Social and behavioral sciences	Government	Assessing social media to assess group threats
Computer science, medicine	Health	Predicting risk of decompensation in complex patients
Social and behavioral sciences, politics, religious studies	Government	How to improve negotiations between groups with different beliefs
Urban planning, architecture	Government, transport, service	What can twitter tell us about how urban populations interact
Computer science (cybersecurity)	Communications, information technology, business	Detecting and minimizing network intrusions
Business, commerce	Finance	Think like a fraudster to prevent credit card fraud

These experiences have led to increased relationships with the client and employment opportunities for the students working on those projects, which can be further developed as the program matures and grows. The practical experience that the MSDS students receive is one hallmark of the MSDS program. Capstones will be expanded as part of the SDS as the MSDS cohort expands both on-Grounds and online.

Presidential Fellows

The Presidential Fellows program is a joint venture between the DSI and the Office of Graduate and Postdoctoral Affairs. Started with support from the Jefferson Trust, this one-year program creates opportunities for graduate students in varying disciplines to work together on collaborative, multi-disciplinary research projects that address real-world problems using

¹⁵ <https://datascience.virginia.edu/projects/capstone>

traditional research methods along with cutting-edge data science tools and techniques. Students are mentored by the PhD supervisors and DSI faculty. Students gain experience in data science techniques which they take back to their own laboratories. Projects for this year are summarized in Table 5.

Table 5 2018-9 Presidential Fellows Research Projects

Discipline(s)	Economic Sector	Title
Systems Engineering, Education	Education	How to turn confrontations in the classroom into teachable moments
Systems engineering, Psychology	Health	Therapist in your pocket
Medicine and computer science	Health	Improving the diagnosis of Parkinson's disease through graphical modelling of brain MRI
Environmental science, urban and environmental planning	Transport, government	Evaluating air pollution exposure and environmental justice by integrating novel high-resolution nitrogen dioxide and human activity datasets
Systems engineering and politics	Government, information	Women and cyber recruitment of extremist groups: A text analytics approach
Psychology, systems engineering	Transport, urban planning	The effect of alarm modalities on drivers' performance in dangerous scenarios

Projects and students for the 2019-20 academic year are being sought. Subsequently we expect to morph this program into the SDS PhD program. The Provost has approved, in principle, the development of a DSI PhD program which will likely be brought before the faculty senate in the second quarter of 2019. Details of the PhD program are described subsequently.

Faculty Research

Table 6 outlines the current funded or proposed research being undertaken by DSI faculty.

Table 6: Examples of DSI Faculty Research Grants

DSI Faculty	Name of Proposal	Agency	Duration	Total Amount
Brown & Bourne	The integrated Translational Health Research Institute of	NIH	3/19-2/24	\$22.6M

	Virginia (iTHRIV): Using Data to Improve Health			
Bourne	Bioinformatics Resource Center for Infectious Disease	NIAID	10/19 - 9/24 (pending)	\$20M
Bourne	Targeting drug resistant bacteria with multi-scale computational models	NIH	4/19-3/24 (pending)	\$2.2M
Bourne	Data for the Public Good; Matching Community Needs to Volunteers with Data Science Skills	Jefferson Trust	3/1/2019-2/28/20	\$50K
Mietchen & Rasberry	Robustifying Scholia: paving the way for knowledge discovery and research assessment through Wikidata	Sloan Foundation	4/19-7/20 (pending)	\$500K
Clark	Argumentation and linked-metadata services for reproducible target validation	NIH	9/16-6/19	\$1.6M
Veeraragavan & Clark	SCH: INT: Waze for cancer: Helping cancer patients and oncologists learn from EHR data	NSF	8/19-8/24 (pending)	\$1.2M
Alvarado	Digital Yoknapatawhpa	NEH		
Flower	Predictive informatics Modeling in the NICU	NICHD	7/10/2014-7/31/2022	
Papin, Brown, Loughran, Skadron	Big Data 2 Knowledge	NIH	8/1/16-7/31/20	\$1.43M
Veeraraghavan, Brown, Davidson	P-CORE: Privacy-Enhanced Coordinated Enterprise Defense via Temporal and Topological Representation Learning	DARPA	7/1/18-6/30/21	\$7.75M
Brown	Design of a Culturally Aware Chatbot Interface	ARO	9/1/18-8/31/19	\$150K
Warren, Brown, Patterson, Saathoff	The Creation Of Muharjirat in America: Social Media as a Platform for Crafting Gender-Specific Interventions for the	NIJ	1/1/17-12/31/19	\$780K

	Domestic Radicalization of Women			
Beling, Brown, Scherer	Center for Visual and Decision Informatics (CVDI)	NSF	10/1/16-9/30/21	\$500K
Brown	Test and Evaluation of Performative Approaches to Predicting Violence by Value-Based Groups	ARO	8/11/18-9/30/19	\$100K
Brown	Hyperparameter Optimization for Object and Event Detection from Patterns using Deep Learning	NPS	1/1/19-8/31/21	\$450K
Rasberry	Establish and share journalism source metadata with Wikimedia	Knight	3/19-3/20 (proposed)	100K
Heger, Jeschke, Kraker, Mietchen	Towards an open, zoomable atlas for invasion science and beyond	Volkswagen Foundation	6/19-6/22	500K
Rasberry	Foundations of Wikipedia health publishing in India	Gates	4/19-4/20 (proposed)	340K
Anderson	Virginia Open Data and Technology Course Curriculum	Commonwealth Provost's Initiative	12/18-12/19	\$50K

These are grants awarded to / submitted by DSI faculty with data-science emphasis, but processed through other schools. The above research via capstones, presidential fellows and DSI research faculty does not fully express the interdisciplinary nature of the research being undertaken. Table 7 outlines examples of what has been achieved with the DSI across some schools, the library and with partners in northern Virginia. It is only a subset as all schools are in some way touched by research within the DSI. A model to be expanded as part of the SDS such that the scope of interdisciplinary engagement will touch every one of the 11 schools, the pan-Grounds institutes, the library, Wise, and northern Virginia.

The 5 Pillars of Data Science	Example Collaborations Within Some Schools, the Library & Northern Virginia									
	Curry (Satellite)	SOM (Satellite)	College (Satellite)	SEAS	Darden	Library	NOVA	All Schools		
Data Acquisition	Biosensors - diabetes mental health etc.	Jefferson Labs	Open hardware lab	Joint MSDS/MBA Joint PhD planned Analysis of financial markets Credit card fraud Customer satisfaction	Wikidata & Wikipedia	Inova pediatric data, USG engagement	Student internships within the region	Presidential Fellows Capstones		
Data Engineering					Access to open science					
Analytics	Precision Education Autism	Translational research (THRIV)	Assessing extremist threats, Assessment of fake news		Faculty profiles, bibliometrics	Genomic medicine				
Visualization & Dissemination										
Policy, Law & Ethics	Data governance	Center for Ethics & Justice	What do you want vs what do you need		Open access policy	Science for the public good				

Table 7 Example Research Activities in a School Without Walls

Guiding Principles

Using this experience of 4+ years of DSI research, the guiding principles that define the SDS research strategy can be stated as follows:

- The research strategy must support the SDS mission, notably as it relates to openness, diversity, and interdisciplinary collaboration.
- The research strategy must be closely aligned with the UVA strategic plan to be announced in June 2019.
- Demand presents an enormous number of research possibilities. The SDS should have a balanced research portfolio that enables collaboration across Grounds with all disciplines.
- SDS research should not compete with similar research being conducted elsewhere at UVA, but instead focus on facilitating large, interdisciplinary proposals that bring together researchers from across the University.
- Our faculty hiring will in some ways dictate the research agenda or, more importantly, we should hire to meet a research (and teaching) agenda. By analyzing the research landscape, we can determine the areas for key hires to complement current strengths and build up areas of determined strategic importance.

General Features

So what should our research strategy be in general terms? We should have active research programs across our five pillars of data science (embodied in Centers as described subsequently):

- Data acquisition

- Data engineering
- Data analysis
- Data visualization and dissemination
- Ethics policy and law

Such research will be sustained by five centers led by tenured and likely endowed faculty which are discussed below as part of the affiliations with other campus units.

Key Areas

Beyond these pillars of data science, we have identified some key areas where we want to have impact based on societal needs. Others will emerge as the SDS grows and new faculty with recognized skills are brought on board. These are some of the areas identified thus far.

Biomedical Data Sciences

The DSI and its faculty were instrumental in the success of the \$22.6M integrated Translational Health Research Institute of Virginia (iTHRIV) proposal, which is a strong indicator of a successful academic medical center in moving research from the bench to the bedside. Data science for personalized medicine is a hallmark of iTHRIV, as is the concept of collaborative team science. The DSI facilitated the movement of the Biocomplexity Institute from Virginia Tech to UVA and DSI faculty are involved in the Biomedical Data Sciences cluster hiring. These are important activities in building a complement of researchers in a field which will be heavily influenced by the vast amount of genotype to phenotype to population data that are being accrued - so called precision medicine. The DSI has joint faculty with Public Health Sciences working on Alzheimer's research and is deeply engaged with the Center for Diabetes Technology and with predictive modeling in cardiovascular medicine and pediatric medicine.

As an illustration of emerging cooperation with the School of Medicine, DSI is helping establish a MS program in biotechnology and a specialization in computational biology as part of the future SDS PhD program.

Educational Analytics

SDS will work closely with Bob Pianta, Dean of the Curry School of Education and Human Development, to establish a satellite (see below) supporting educational analytics. As we gather more data on students and teachers, we can provide more targeted interventions to improve the quality of individual education and teacher quality. We refer to this as precision education, which nicely illustrates the potential value of the SDS. Since the SDS is also engaged in precision medicine, we expect synergy between these two formally discrete fields. Targeting the individual -- whether for medical or educational improvement -- has common underpinnings ranging from how data are collected to the privacy concerns and consent issues that must be addressed. Through the SDS, there is the opportunity to collect and use data to implement changes and interventions, possibly in real-time, to help students while maintaining best practices for privacy and security.

Cybersecurity

The DSI has been a major contributor to the creation of a cybersecurity research program at UVA. Through the efforts of the DSI, the UVA network is the basis for advanced analytical work in the identification and response to malicious software and hostile attacks. This has resulted in a major award from DARPA.

Democracy

The College of Arts and Sciences has established Democracy Labs which have parallels to the SDS in that they are both highly interdisciplinary, engaging faculty from across Grounds in joint research and teaching initiatives. The DSI Director, A&S Dean, and Democracy Lab Director have been exploring a shared faculty appointment as well as shared capstone projects by DSI students.

Environment

The DSI has a history of working with the Department of Environmental Science and more recently the Resilience Institute through joint research projects, sponsored seminars and workshops. Given the societal impact of a changing environment, the SDS will pursue a joint faculty hire in some aspect of environmental analytics as yet to be determined.

Business, Commerce and Finance

The DSI has had a good relationship with the finance industry, notably with Capital One and S&P Global, where MSDS students have repeatedly undertaken capstone projects exploring areas of societal concern such as detecting credit-card fraud and enhancing customer satisfaction. Consequently, these companies have hired a number of our students. In business and commerce, the DSI offers a dual-degree program with Darden. A business concentration as part of the proposed SDS PhD program is being discussed with Darden. Recent discussions with Darden and McIntire have been around a satellite (see below) with an initial focus on executive education. The DSI hired a research fellow to build out this research area which is currently focused on time series data associated with financial transactions.

Affiliations with other campus units envisioned?

School of Data Science Satellites and Centers

The School of Data Science (SDS) is being established as a school without walls. Already partially in place, the SDS intends to support a network of collaborative interactions within and across all existing University of Virginia (UVA) schools, including Wise, northern Virginia, and the administration. This is illustrated in Figure 4.

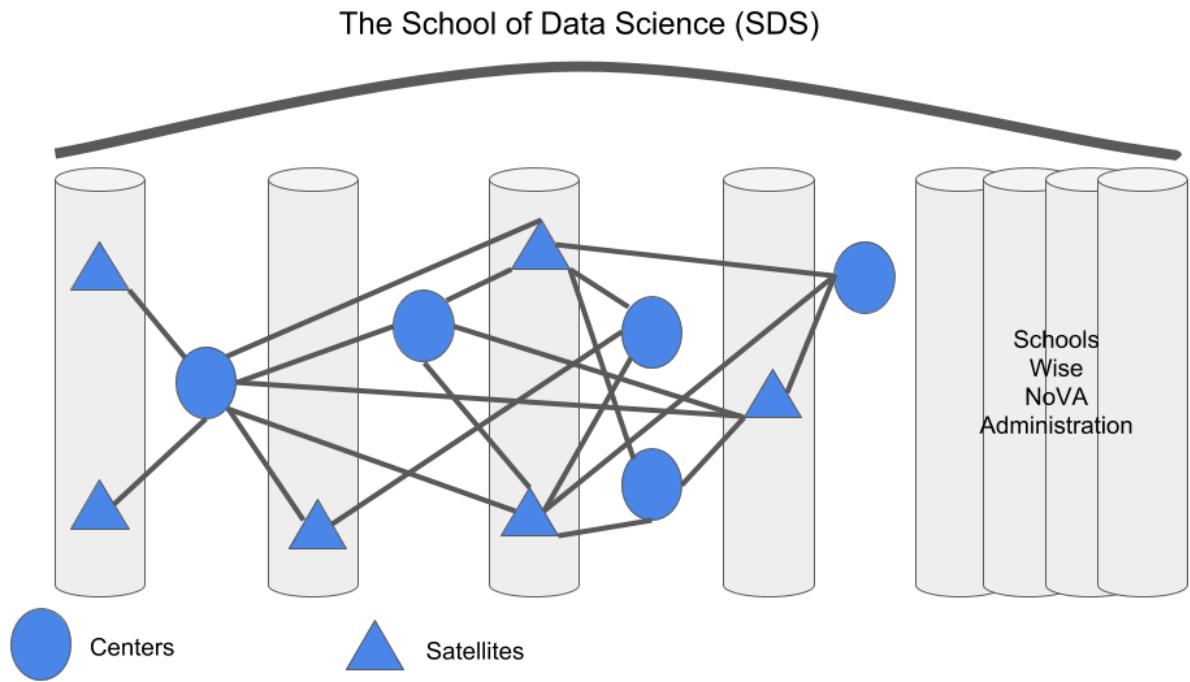


Figure 4 Relationships Between SDS Satellites, SDS Centers and UVA Writ Large.

Affiliations will take place using a center and satellite model. Satellites are typically embedded in other Schools; centers are analogous to departments.

Centers

Centers replace traditional departments within the SDS with the goal of making them dynamic, agile and more easily phased in and out as the needs of data science change.

Centers bring together faculty, staff, and students (stakeholders) who share a common interest. Features of a center are as follows:

- Each center will have a center director who is anticipated to be a tenured faculty member and ultimately an endowed chair.
- Center directors report to the SDS Dean.
- Center activities are overseen by the Office of Center and Satellite Affairs.
- Stakeholders can be part of more than one Center.
- Centers will be established for a period of 7 years.
- Centers will be reviewed after 5 years by a review committee from across Grounds and beyond and a recommendation made to the SDS Dean to either terminate, continue, or change direction for a further 7 years.
- If a Center's status changes, stakeholders will become part of one or more other (or modified) Centers.
- The formation of a new Center will be done in discussions between the SDS and other Deans and recommended to the Provost for review and approval.
- Centers have faculty whose primary appointment is with the SDS, not the Center.

- Centers have affiliated faculty with primary appointments in other schools, offering immediate and ongoing collaboration opportunities with existing faculty in other schools.
- Affiliated faculty will have the opportunity to join a different Center if its status changes.

A variety of centers exist or are planned, starting with what we regard as support for the pillars of data science (Figure 5):

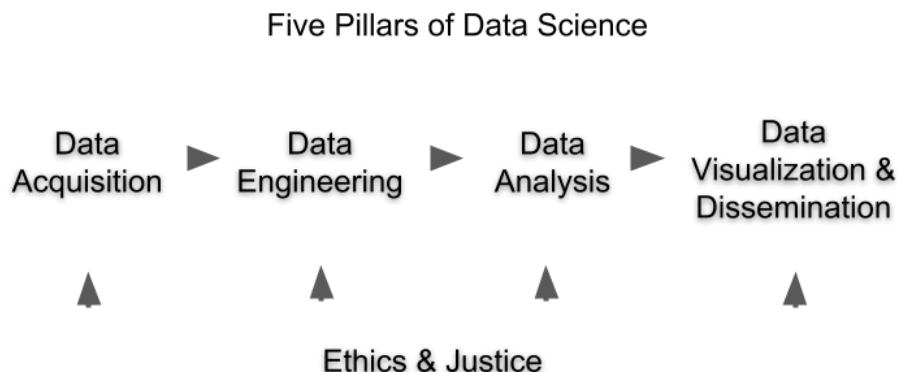


Figure 5 The SDS Pillars of Data Science

- **The Center for Data Acquisition and Analysis** works with sensors, Internet of Things (IoT), and other devices to assemble and analyze a large variety of data. Examples include, apps that collect data from smartphones, biosensors such as the artificial pancreas, GPS tracking devices, sensors collecting environmental data, satellite data collection and analysis. Maker spaces that allow students to design and assemble sensors from inexpensive materials.
- **The Center for Data Engineering** specializes in data representation, data quality and security. Examples include, cybersecurity, privacy, ontologies, interoperability, and data fusion, etc., to make data useful for a broad community while preserving the privacy rights of individuals.
- **The Center for Data Analysis** specializes in the application to different forms of machine learning, such as deep neural nets and natural language processing.
- **The Center for Data Visualization and Dissemination** specializes in making data and analytics output as usable and interpretable as possible. Examples include expanding the [open data lab](#), an initiative with the library for accessibility to reference data and associated tools which are available to a broad community, building out resources and tools in data visualization, and specializing in alternative representations of data such as sound.
- **The Center for Ethics and Justice** cuts across all activities. Increasingly, anxieties and concerns expressed by the public, private companies, and governments indicate a need for such an initiative. The Center for Data Ethics and Justice will respond to this need and foster a commitment to ethical reflection and imagination regarding the most pressing challenges of a data-centric society through research, education and policy/community engagement. From the challenges of bias, discrimination and inequality, to the threat to democratic values and practice, to questions raised by the

relation between humans and AI technology, to those of openness and accountability, the Center for Data Ethics and Justice will foster research and education that critically engages these pressing challenges, with the intent of offering alternative and innovative ethical responses to help build a better world for all.

Beyond the pillars, centers in sports analytics, materials analytics, and financial analytics are examples of other centers that may be established over time. Each center will be led by tenured faculty member to provide stability who is capable of building out a major research and training initiative. Proposals for centers will be vetted by the Academic Executive Board (which includes representation from across Grounds as described subsequently). The formal process of creating a Center is already established by the University with approval sought through the Provost's office. The Center for Ethics and Justice was created this way. The Provost and Provost-elect are supportive of the Center model as presented here. Centers and satellites form a network of collaborative interactions across Grounds.

Satellites

SDS satellites exist specifically to provide a focus of collaboration within a specific school, the College at Wise, and northern Virginia. Example satellites that the DSI is currently putting in place include a satellite in Curry to enable educational analytics; in the School of Medicine (SoM) to enable biomedical data sciences; a satellite in the College (an extension of the DSI Director's laboratory) to provide interactions in biology, chemistry, and biochemistry; and a satellite in northern Virginia to support executive education. Satellites will be led by tenured faculty members with joint appointments in the SDS and (if a school) the school where the satellite is located. A satellite does not necessarily imply a single physical location or domain area. For example, the emerging satellite in SoM will occupy space in the Fontaine Research Park, Public Health Sciences, and in temporary space allocated to the proposed SDS on the corner above the Ragged Mountain Athletics store (described subsequently).

Activities of an SDS satellite located in a school include, but are not limited to:

- Working space for jointly appointed faculty, staff and students.
- Shared responsibility for data science related seminars, courses, tutorials and workshops pertinent to that school.
- Advising students and consulting with researchers on data science related matters.
- Active exchange of best practices between the satellite, the SDS, and other satellites and centers.
- Joint grant development, creating more opportunities for funding through strong ongoing collaboration across disciplines.
- Assuring communication and collaboration between the SDS and the respective school
- Research between SDS faculty and students and those of the school.
- Focal point to explore data governance issues related to the specific disciplines represented by the school.
- Longer term: possible maker space (in collaboration with Biomedical Engineering and others).

Consider two examples of the interrelationships between satellites and centers.

Example 1 - A geographically strategic satellite. There is enormous opportunity to expand UVA's reach into northern Virginia (NoVA) for research and educational opportunities. SDS is looking at leasing space in the Rosslyn facility already occupied by Darden and the Biocomplexity Institute. This will foster interrelationships between these entities, but also provide a focal point to offer executive continuing-education courses in data science and computer science and related activities in collaboration with SCPS, SEAS, and other interested schools. To provide these courses the NoVA satellite will draw on expertise from other satellites and centers.

Example 2 - A Center for Data Acquisition and Analysis. Data are being acquired across all schools, often using similar technologies (e.g. cheap sensors and Raspberry Pi's), but no focal point exists to exchange best practices. Similarly, an enormous amount of research has been done on the artificial pancreas (a biosensor) that could benefit other emergent data-acquisition projects. Data acquired from a variety of geospatial projects related to the environment, smart cities, reconstruction of historic sites, etc., occur in isolation. The Center for Data Acquisition and Analysis, through interaction with faculty in the Link Lab, SDS satellites and beyond, will have a forum through which to collaborate and exchange best practices, mentor students jointly, etc.

Interactions with other Pan-University Initiatives

Appendix C provides letters of support from existing pan-University initiatives with which the DSI is already collaborating. These letters endorse expanding these activities. The following are examples of activities under way:

- **Library** - The DSI collaborates with the library on a number of training and research projects, and this will continue with the SDS. Examples are:
 - Scholars Lab - working with Allison Booth on various digital humanities projects.
 - Wikipedia - Trust and safety, Wikimedia and Cochrane - expanding UVA's presence
 - Wikidata - Scholia - a bibliometrics project
- **Environmental Resilience Institute** - The DSI has conducted capstones and presidential fellowships in environmental science.
- **Brain Institute**
 - Currently, an existing DSI faculty member is engaging with the Brain Institute to establish a seed grant for the purposes of studying the relationship between arterial blood pressure and oxygen perfusion to the brain in infants with hypoxic ischemic encephalopathy undergoing brain cooling therapy using NIRS and aEEG technology acquired through the proposed grant. In addition, the autoregulation of very low birthweight (VLBW) preterm neonates with intraventricular hemorrhage (IVH) will be studied.
- **Global Infectious Disease Institute**

- A Capstone project sponsored by Christopher Moore, M.D. of the Global Infectious Disease Institute investigated the early detection of sepsis in intensive care unit (ICU) patients.
- **Institute for Advanced Technology in the Humanities**
 - Collective Biographies of Women project- initially annotation of the corpus using natural language processing. Longer term goal to understand women and their representative role in society at that time.
 - Curry Center for Advanced Study of Teaching and Learning (CASTL) Real-world human interactions involve multiple brains dynamically engaging together. However, traditional neural metrics consider individual people in isolation. This project focuses on developing a metric for quantifying and optimizing cognitive performance and engagement using simultaneous recording and integration of multiple streams of neural data from student-teacher pairs interacting with each other.
- **Center for Advanced Medical Analytics** - When acutely ill patients are admitted to the hospital, clinicians determine diagnoses and appropriate therapies for these patients with the aim to restore or improve health. Despite best efforts, deterioration of patient condition and consequent escalation of care requiring transfer to an Intensive Care Unit (ICU) occurs at a rate of four to five for every 100 acute care admissions. Since mortality increases with every hour of delay in transferring critically ill patients to the ICU, the healthcare community has a vested interest in identifying patients at risk for deterioration, and subsequent potentially catastrophic conditions, as early as possible.

Education - Curriculum

THE SENATE IS NOT BEING ASKED TO APPROVE ANY NEW PROGRAMS AT THIS TIME.

New programs will come through submission by the SDS to the usual Senate approval process at a later date (see timeline).

The vision for educational programs offered through the SDS would include an expansion of the already existing (now in its 5th year) DSI Masters in Data Science (MSDS) program into undergraduate, terminal degree, and professional programs. In all programs we have, or will develop, a defined set of core competencies, and we will bring together or establish courses to meet those competencies. In addition, courses offered throughout the University taken as electives provide the means for students to gain experience in a given practical area of study. A hallmark of all programs will be an emphasis on responsible data science and practical application through undergraduate research projects, MSDS capstones, and full-fledged PhD research projects, as described below.

The development of new programs will be the responsibility of the SDS Academic Affairs Committee (Fig. 6) led by a tenured faculty member which, as described, will comprise members from the SDS team (faculty and staff) as well as representatives from across Grounds, notably from the key departments that have been identified. The SDS Academic Affairs Committee expands the DSI Academic Affairs Committee and processes to have greater

representation from across Grounds and employs expertise in undergraduate, on-line and executive education in developing and supporting the educational program introduced.

Vision for degree offerings: names and descriptions

The vision for the degrees the SDS would offer and number of students anticipated were outlined in Table 1. Numbers are estimated from a combination of market forces, comparisons of enrollments at other universities, a desire to maintain a high-quality program producing data-science leaders, a sensitivity to impact on other UVA departments and schools, and quotas set by the Provost's office. **Final enrollment is not defined by the SDS**, but will be set within the confines stated by the Board of Visitors and related through the Provost's Office to the SDS Academic Affairs committee (Fig. 6), with representation from key departments whose programs and courses might be impacted.

Here we introduce each proposed and existing degree to provide the needed context for understanding the SDS as proposed. Development of new degrees will be led by a tenured faculty member.

Undergraduate certificate in data science

A committee with all schools represented and chaired by the Vice Provost for Academic Affairs and the DSI Director is addressing this issue. The certificate is anticipated to consist of 12 credits. Thus far, the committee has defined a set of core competencies, identified the courses offered across Grounds that meet those competencies, and identified the gaps that exist. The objective is to avoid uncoordinated efforts for certificates offered by all schools, by offering a smaller coordinated set of certificate tracks, e.g., STEM, non-STEM. In this way a certificate in Data Science means something similar across Grounds but accommodates for broad differences in student interest. The administration of the certificate across schools has not yet been determined. The committee hopes to present a proposal for an undergraduate certificate in data science to the Faculty Senate in the second quarter of 2019.

Undergraduate minor in data science

This has yet to be developed, but it is anticipated that it will be developed in the same way as the undergraduate certificate, that is, with representation from all schools, but comprising of additional prerequisites, course credits, and project work in alignment with the University standards for an undergraduate minor. It is anticipated that a student would take the certificate or the minor, but not both. Again, this will be developed by consensus from across the Schools and come before the Faculty Senate for review at a later date.

Undergraduate degree in data science

This has yet to be developed and is currently envisaged as a 4-year program that attracts its own student body rather than siphoning from existing pools of students in other schools. It is anticipated the curriculum would be interdisciplinary by leveraging a number of existing courses within other schools. This would contribute to the intellectual and financial wellbeing of all schools, since 75% of the revenue would flow to the school offering the course.

Masters degree in data science (Grounds)

The MSDS is a well-established and successful 11-month degree program run by the DSI. Details can be found on the DSI website¹⁶. The MSDS is also offered as part of a dual-degree program with an MBA (Darden School of Business) or MD (SoM). The MSDS and dual-degree programs will be expanded both in numbers and in scope. In terms of numbers, the intent is to expand to a total of three 50-student cohorts from the current single program of 69 students. In terms of scope, the intent is to offer dual degrees such that PhD programs from any school could offer the opportunity for their students to complete an MSDS before, after, or during their PhD training. There is a great interest in collaboration with other schools to continue to develop dual-degree graduate programs that embrace the established MSDS. The MSDS program emphasizes and will continue to emphasize, practical training and responsible data science. Employment statistics are described in Figure 2.

Masters degree in data science (Online)

The MSDS online has been under development for almost a year and began with the selection of a third party that could help design, market, and launch an online MSDS. After a procurement process, Noodle Partners was selected as the vendor; admissions are open and classes will begin in summer 2019. Noodle Partners was selected in part because it offers a platform that can be integrated with our existing student-management systems. The software, processes, and content developed for the platform will be owned by UVA, and the platform and experience could be used by other schools, highlighting SDS's desire to enable all of UVA.

PhD program

The DSI hired an Associate Director for Research Development in late summer of 2018, and she has been charged with preparing a PhD program submission, first to the Faculty Senate and later to SCHEV. We expect to present a proposal to the Senate for the PhD program in the second quarter of 2019. While details could change, the current plan under discussion is as follows.

The proposal will include a list of prerequisites, core courses, and an initial list of electives that form the course requirements for the PhD program. Students admitted to the program will be assigned an SDS faculty mentor based on their interests and those of the mentor. Students will be expected to undertake 2-3 rotations with UVA faculty or with corporate partners. At the end of those rotations, students will take a qualifying exam to determine if they have the potential to succeed as independent researchers in data science. Having qualified, from these research experiences or others under advisement from the SDS faculty mentor, the student will select a PhD project. As such, the students will have dual mentors, the SDS faculty member and either a faculty member from another school or a private-sector mentor. Much of the PhD student funding will come from the SDS Student Fellowship fund, a \$40M endowment as part of the gift to SDS. As part of our commitment to societal benefit, SDS hopes to leverage the existing Center for Ethics and Justice to support PhD students specifically studying ethics and justice

¹⁶ <https://datascience.virginia.edu/degrees/msds-degree>

with co-mentors potentially in Batten (public policy), School of Medicine (bioethics), or A&S (democracy), and possibly others.

Professional certificate in data science

Thus far, the DSI has not offered a professional certificate; however it has offered one-week boot camps for professionals and has identified that a significant market exists for such a program, particularly in northern Virginia. At this point, we propose that the SDS work with the School of Continuing and Professional Studies (SCPS) to map out what a joint program would look like with a view to offering such a program in northern Virginia. The DSI has affiliated faculty in northern Virginia who could be instrumental in taking this program forward. The SDS will also consider collaborative arrangements with companies that are interested in integrating their customized training with data science skills in order to maximize the utilization of courses/modules developed for the executive-education format.

Research experiences for undergraduates

Recognizing that students at the undergraduate level are increasingly capable of data science research, the DSI applied for a NSF Research Education for Undergraduates (REU) grant with a focus on minority students. This was not awarded for 2019, but the DSI was encouraged to re-apply. Doing so under the SDS banner will add credibility to a proposal for a 12-week summer experience for students. In the interim, the DSI will offer such an experience in the summer of 2019, thanks to a generous gift from one of the DSI Board members.

Accreditation

This has been discussed with the Director of the Office of Institutional Assessment and Studies. The SDS will be in compliance with the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC) accreditation standards.

IV. Faculty & Staff

The school shall have a defined faculty & staff sufficient to fulfill the school's mission, goals, and objectives.

Overview

Any school is only as good as its human capital, its faculty, staff, and students. Over time high-quality, career-satisfied faculty and staff will create the best research and teaching environment and the most satisfied students. This section describes what is in place already through the DSI, the types of faculty and staff to be hired, their roles, and how endowments will be distributed.

What the DSI has in Place

The DSI treats tenured faculty, tenure-track faculty, general faculty, and staff as *team members* who will all have some say in the operation of the DSI. This has worked well for the DSI which, while small (19 FTEs), has gained from everyone having a voice in its operation and goals. This

is reflected in the Provost-approved DSI Faculty Handbook¹⁷ which was developed with reference to that of other UVA schools. That handbook will be expanded and adopted, with necessary modifications, by the SDS after approval from the Provost's office. This is referred to subsequently as the SDS Faculty Handbook.

Within the DSI morale is very high, we believe as a result of this team-oriented approach with everyone pulling together to achieve the mission. Indeed, the success of this model and the commitment from the team as a result of it, are, in no small part, how we have reached this point of SDS formation, which includes administrative and leadership endorsement and securing the largest gift in the history of UVA.

The DSI currently consists of tenured, tenure-track (in other schools), general faculty, and staff as follows:

- Director (tenured faculty, SOM)
- Founding Director (tenured faculty, SEAS)
- Associate Director for Operations and Strategic Initiatives (staff)
- Associate Director for Executive and Continuing Education (staff)
- Associate Director for Research Development (staff)
- Associate Director for Career Connections and Community Engagement (0.5FTE; staff)
- Program Director MSDS (staff)
- Assistant Director for Career Advising, Data Science & Analytics (0.5FTE; staff)
- Assistant Director for Academic Affairs (staff)
- Data Scientists (3; PRS faculty, staff)
- Ethicists (1.5; tenured faculty, A&S & PRS faculty)
- Communications manager (staff)
- Admissions and Recruiting Coordinator (staff)
- Office manager (staff)
- Event Planner (staff)
- Wikimedian in Residence (staff)
- Writer (part time student)

The team of 19 DSI professionals is supported by a number of tenured, tenure-track, and general faculty from other schools¹⁸ who teach in the MSDS program. Thus far, staffing has followed a step function; only when there are resources to support an expansion of activities do we make new hires. This model of sustainability will continue as we move to become the SDS. Initially, each DSI team member will assume the same role in the SDS. The Provost's office has provided valuable advice on staffing in moving from the DSI to the SDS.

¹⁷ <https://drive.google.com/open?id=1ktEEwMDEcJOKeazZuxfln28g1da9bEZM> - approved by the Provost's office.

¹⁸ <https://datascience.virginia.edu/people/faculty>

Faculty and Staff Going Forward

However well the SDS operates as a team internally, there must, of course, be a mapping to HR classifications including titles that have meaning across the institution. This is important for establishing joint appointments and a sense of equity across the institution. To this end, the SDS will support tenured faculty, tenure-track faculty, general faculty, and administrative staff.

Tenured Faculty

Tenured faculty have promotion and tenure in accordance with the SDS Faculty Handbook, which is in alignment with tenure in other schools but also expresses SDS guiding principles, including openness, value of research products beyond papers, and evidence of interdisciplinary collaboration. These faculty will play the following roles in the management and development of SDS in addition to their individual research, teaching, and administrative responsibilities:

- Recommendations to the Provost for School leadership
- Defining strategic directions - including new research programs
- Design and execution of degree, certificate and executive education programs
- Leadership of SDS Centers and Satellites (as described below)
- Hiring of new faculty
- Oversight of promotion and tenure of existing faculty
- Support of the SDS Guiding Principles

Tenure-Track Faculty

Tenure-track faculty have promotion and tenure review in accordance with the SDS Faculty Handbook, which is in alignment with promotion and tenure review in other schools, but also expresses SDS guiding principles, including openness, value of research products beyond papers, and evidence of interdisciplinary collaboration. Until they achieve tenure, tenure-track faculty will play a lesser role in the management and development of SDS, focusing more on their individual research, teaching, and community service.

Tenure and Tenure-Track Appointments

The SDS will make a number of tenured and tenure-track appointments at 100% within SDS. However, given the interdisciplinary nature of data science, we anticipate a significant number of tenured and tenure-track appointments to be dual appointments with other schools.

We expect that the faculty leading the SDS Centers will be 100% SDS tenured appointments, in addition to 100% SDS tenured appointments that are not leads of centers. We envisage faculty leading satellites (described below) to be jointly tenured appointments as discussed below. Tenured appointments provide continuity and recognition in key areas of data science that will be expanded as the SDS grows.

Joint Appointments

Joint appointments are already in place at the DSI and achieved through the use of Memoranda of Understanding (MOU) between the DSI and the department where a faculty member has his or her tenured or tenure-track appointment. What will change under the SDS is that the faculty member may have part of their tenure within SDS. Joint appointments now typically have a primary (the major percentage of effort) and secondary department (the minor percentage of effort), perhaps within different schools. The teaching and research commitment to each department is clearly defined when the faculty member is hired, and rules are in place to change that distribution when department chairs agree it makes sense to do so. This model will also be used by the SDS. However, we also propose to promote true interdisciplinarity through 50:50 tenure and tenure-track appointments.

If the appointment is 50% within SDS and 50% within another school, then a promotion-and-tenure committee will be formed with members from the SDS and the other school to provide a truly interdisciplinary evaluation. This instills a sense of belonging by the faculty member in both the SDS *and* the other department, a needed characteristic for true interdisciplinarity. Note that a 50% time commitment in a department outside of SDS could imply a *more than* 50% financial commitment from the SDS. The SDS will use its resources judiciously to seed important areas of interdisciplinary activity with the belief that, over time, this will lead to a greater return through grants received and students trained than would be the case if the faculty member were located indeed solely within the SDS. All that is stated here regarding joint appointments is consistent with UVA policy¹⁹. 50:50 appointments are anticipated for those faculty leading SDS satellites in other schools.

General Faculty

The model for general faculty is already in place and operational within the DSI. General faculty have a designated amount of time assigned to teaching, research, and service as defined in the SDS Faculty Handbook. General faculty also have a defined evaluation and promotion process (outlined in the SDS Faculty Handbook) in alignment with the University's Academic General Faculty policy.

General faculty will play the following roles in the management and development of SDS:

- Undertake teaching in SDS (now DSI) mnemonic courses. Teaching loads will vary depending on the nature of the appointment. A typical 9-month full-time teaching appointment would be 3-3; a twelve month full teaching appointment would be 3-3-2.
- Undertake research, including as PI's and named investigators on grants and contracts in accordance with the nature of their appointment.
- Work with tenured faculty to develop and support SDS centers and satellites.

¹⁹ <https://uvapolicy.virginia.edu/policy/prov-029>

- Contribute to the service function of the SDS, for example, serving on committees and representing the SDS at career and promotional events.

Staff

The pro forma budget describes additional staffing beyond what the DSI already has in place. It is difficult to determine exact needs/job functions at this time, but we will remain within the budget outlined. Staff affairs will propose and oversee new staff hiring based on their own initiative and from a strategic perspective based on input from the Advisory Board, the Responsibility Diversity, Accessibility and Transparency group and the Academic Executive Board (described below)

New proposals for staffing changes impacting SDS function will be posted internally for review and comment by faculty and staff. After consideration of any comments, appropriate functional units will be directed to make the appropriate hires by the Dean or designee.

Future Numbers and Growth

Because faculty and student numbers in SDS are subject to approval by the Provost and BOV, the absolute numbers and the faculty:student ratio are unknown. Given these unknowns, our goal remains maintaining a faculty: student ratio of 1:15 in accordance to a number of the other UVA schools.

In terms of the distribution of tenured, tenure-track, general faculty, and staff we have established the estimates outlines in Table 8

Table 8a: Estimated Faculty and Student Numbers within the SDS

	19-20	20-21	21-22	22-23	23-24
Undergraduate Certificate		100	200	400	700
Total Undergraduates			25	50	100
Total Graduates (MS+PhD)	85	199	338	362	366
Total Executive Education	30	60	90	120	120
Total Students	115	259	453	532	586
Tenured & Tenure-track Faculty	10	17	23	27	27
General Faculty	6	10	12	12	12
Total Faculty	16	27	35	39	39

~Faculty : Student Ratio	1:7	1:10	1:13	1:14	1:15
Staff	14	16	18	20	20

Table 8b: Estimated Student Numbers outside of SDS

	19-20	20-21	21-22	22-23	23-24
Undergraduate Certificate		100	200	400	700
Undergraduate Minor		50	100	150	175

Notes on Tables 8a:

- Re students: Since undergraduate certificates will involve courses taught in other schools (as best determined at this time) this does not figure into the Faculty : Student ratio.
- Re faculty: This is the total number of FTE's dedicated to SDS and includes buyouts from other departments for specific courses. For example, a faculty member who is appointed 50% in SDS and 50% in another school, would be counted at 0.5 of the 16 total faculty in 2019-20.
- Online materials will be developed mainly by tenured or tenure-track faculty, but in subsequent years delivered by general faculty.
- For sustainability and continuity satellites and centers will be led by tenured faculty.

Notes on Table 8b:

- These numbers are best estimates at this time.

Endowments

Based on the current endowment (see resources), it is anticipated that over time, the SDS will appoint on the order of 10 endowed chairs to tenured faculty allocated across early-, mid-, and late-career levels. The exact distribution will depend on the quality of the candidates. The goal is to hire the absolute best faculty in a competitive marketplace. Like all tenured appointments, endowed chairs will be approved by the Provost and BOV.

The gift and matching from the Bicentennial Professorship Fund create a total of \$63M in professorship endowments. The income will be allocated at the discretion of the SDS Dean in discussions with UVA leadership and the Academic Executive Board (see below). Of the 10 endowed chairs, 5 will support leaders of the 5 centers that support the fundamental pillars of data science, namely:

- Data acquisition - e.g., sensors, imaging devices, crowdsourcing
- Data engineering - e.g., cybersecurity, data standards, knowledge engineering

- Data analysis - e.g., machine learning, natural language processing
- Data visualization and dissemination - e.g., virtual reality, 3D modelling, open science
- Ethics & Justice - e.g., normativity, consent, privacy laws

For continuity, it is expected that these appointments will be 100% within SDS. The remaining five endowed chairs will be joint with other Schools and key departments within those Schools to begin to create the infrastructure of a truly interdisciplinary SDS. Based on the DSI experience of working across Grounds, we have tentatively identified the following areas where endowments might be deployed to take advantage of existing strengths:

- Biomedicine
- Systems
- Humanities
- Finance

Areas where we wish to build our capacity include:

- Educational analytics with Curry
- Business analytics with Darden and McIntire
- Environmental analytics with the College and Environmental Resilience Institute
- Digital humanities and social sciences with the College, Library and the Quantitative Collaborative, among others
- Leadership analytics with Batten and others

Specific hiring areas will be determined on school formation through recommendations of the Academic Executive Board, leading to a truly interdisciplinary foundation of high quality teaching and research faculty. Each hire will be expected to establish a team of professionals, including postdoctoral fellows and students, to build out the enterprise.

To this end, the gift also includes a \$39M endowment for a fund to support graduate students and postdoctoral fellows. An additional \$3M of expendable funds is included in the gift to facilitate recruitment. Taken together, we will provide enormous incentives needed to hire the best in a competitive marketplace. A portion of these funds will be allocated to incoming faculty and the rest to seed research and teaching in strategic areas identified by the Academic Executive and Advisory Boards.

The gift also includes a \$10M endowment for the Quantitative Foundation Visiting Fellows Fund. This fund will enable scholars to visit and undertake research and mentoring through the SDS. It will be used to support faculty sabbaticals both for faculty internal to UVA and from around the globe. It will also be used to support strategic visits to, for example, establish collaborative research initiatives, by graduate students and postdoctoral fellows. All will hopefully become SDS ambassadors as a result and further the reputation and reach of the SDS. Requests to use these funds will be reviewed and approved by Faculty Affairs in consultation with the Academic Executive Board as needed.

Organization

The SDS will be organized in a way that expands upon organizational principles adopted by the DSI. In recognition that certain departments could be particularly impacted by the formation of the SDS they will have a special role to play in organization and governance and are described below. Those departments include statistics, mathematics, engineering systems and computer science.

We propose an organization for the SDS that consists of a series of functional modules (Fig. 6).

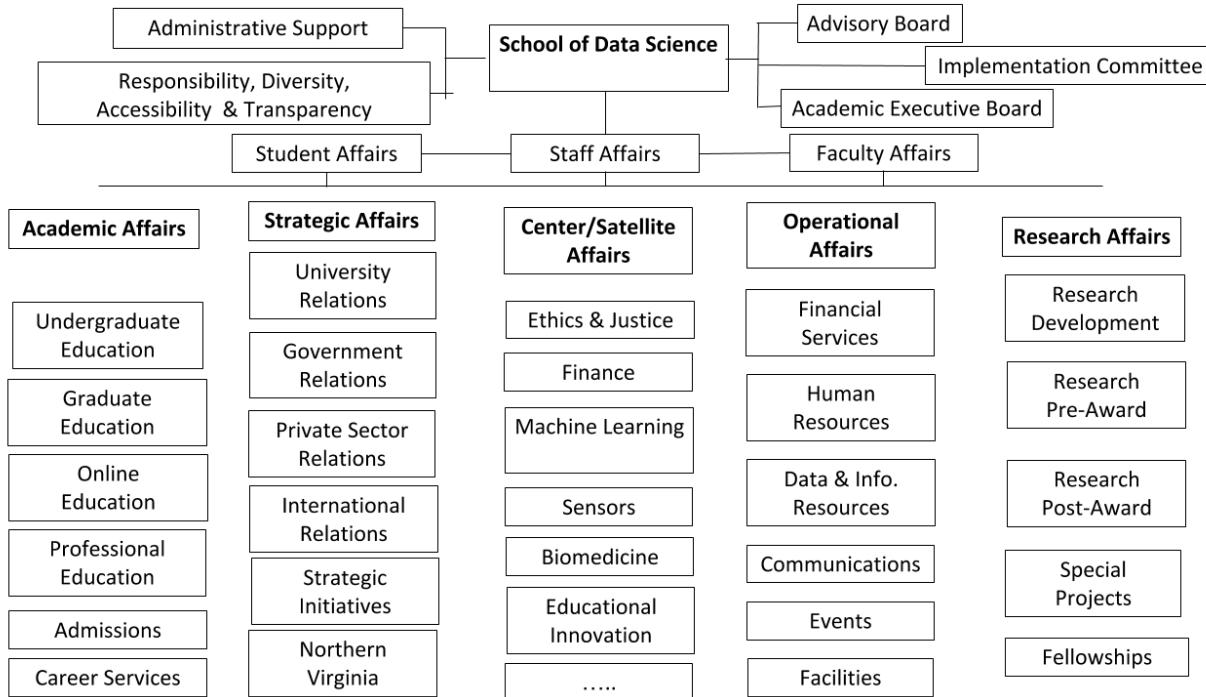


Figure 6 The SDS as a Series of Functional Modules

Note the following organizational features from Figure 6.

- An individual could lead or co-lead more than one functional module. From the viewpoint of the UVA HR system and equity across schools and across institutions, traditional titles will be assigned, for example, Associate or Assistant Dean, but internally, we will use titles that are more group and team oriented as we do for the DSI (see below).
- *The SDS is led by a Dean (TBD)*. Until a Dean is appointed, the Director of the DSI will assume the role of Acting Dean at the request of the President, Provost and Provost-elect.
- *Advisory Board*. There is currently a DSI Board²⁰ of private-sector experts who advise the DSI Director. This group has proven pivotal to the success of the DSI as well as planning for the school and comprises private-sector leaders, some of whom are alumni.

²⁰ <https://datascience.virginia.edu/people/faculty>

The charter²¹ for the DSI Board will be modified to reflect the broader scope represented by the SDS.

- *Academic Executive Board.* This board is comprised of the Provost, Deans, VPR (or their designees), Faculty Senate representative, representatives from key departments, DSI founder, representatives from other academic data-science initiatives, Commonwealth representative, and NGO representatives. A draft charge²² has been prepared for this yet-to-be-formed board. The overall objectives are to assure the SDS's leadership position relative to other institutions and to be sure that the SDS maximizes the integration with other schools and departments across Grounds. The SDS Dean will chair the Academic Executive Board.
- *Implementation Committee.* This committee was established at the behest of the donor foundation and approved by the President to ensure goals and timelines are being met with respect to the gift agreement, which mostly concerns building construction. The committee comprises the SDS Dean, Provost (or designees), VP for Operations (or designee), VP for Advancement, donor, donor's designee, and a representative from the President's office. The committee will meet at least quarterly from January 1, 2019, through December 31, 2021. The first meeting was held February 7, 2019, and the next is April 5, 2019.
- *Responsibility, Diversity, Accessibility, and Transparency.* A proposed pivotal committee with a rotating chair and including the UVA VP and Chief Officer for Diversity and Equity and comprising team members from the SDS as well as diversity leads from other Schools and those across Grounds, including atheist humanists, utilitarians, practitioners of the major religions and philosophers. A major distinguishing feature of the SDS will be a commitment to responsible action around all data science activities. This is crucial at all times and particularly in an era when the misuse of data is prevalent. This is more than ethical conduct when conducting data science, it is also about choosing directions that maximize societal benefit and to determine when data and analytics should be openly and freely available and when not, so as to protect the rights of the individual and beyond. This can only be done in an environment that is inclusive and accessible. This is discussed further below.
- *Student Affairs, Staff Affairs and Faculty Affairs* are committees drawn from those overseeing student, staff, and faculty activities, respectively. There is an overlap of committee members to maximize communication and integration of respective activities. These committees will include members from key departments across Grounds to maximize integration and minimize duplication and competition. Each will be chaired by the Dean or designee who is a tenured faculty member.
- *Affairs* committees are each chaired by an Associate Dean who oversees their activities.
 - Academic Affairs – All aspects of student education and experience from admission application to job security and satisfaction.
 - Strategic Affairs – Visioning and executing on the data science school of the future, for example:

²¹ <https://docs.google.com/document/d/1zaKGWp40qVXU32U37IDCoZPgAWMT5-E-nynPaDZ9F90/edit>

²² <https://docs.google.com/document/d/1w3cFEvQKbYL5Lmidru2L2t1jr-8JQWK1yNa42fJeEfk/edit#heading=h.9lz8piy51ql8>

- Special attention to training and research opportunities in northern Virginia
 - University relations is particularly concerned with the application of data science to the university's own data
- Centers & Satellite Affairs – As stated, the SDS will not have departments. Centers support specific research and educational thrust areas; Satellites support specific disciplines and are found in other Schools or are geographically remote (e.g., northern Virginia, Wise). These activities are overseen by Center and Satellite affairs.
- Operational Affairs - responsible for the successful operation of the School. These functions are mostly self-explanatory and most operational within the DSI framework. The Data and Information Resources will be a special functional module to work with the library and IT services to cooperatively make available data, analytics, information etc. both needed by the SDS team (faculty, staff, students) and offered by the team.
- Research Affairs - Responsible for grants, fellowships and research engagements with the private sector and others. This is largely operational at this time as part of the DSI and will be expanded to meet the needs of the SDS.
- *Team members* (faculty and staff) - Typically belong to multiple functional modules, with each module having a lead. Team leads will liaise with members of key departments across Grounds to maximize integration and cooperativity.

Use Cases

Consider how this will work in practice with four hypothetical (but eventually likely) use cases, i.e. practical scenarios applying this model to the life of the University and the SDS.

Case 1: Education. A new graduate course is proposed in responsible digital humanities to be given both online and on-Grounds. Most likely, this would be initiated by a faculty member, appointed at least partially in the SDS, versed in the field, and likely a member of the Center for Ethics and Justice. The initiator would consult with other Center team members to determine viability and overlap both within and outside the SDS. Assuming there was informal consensus to proceed, the initiator would consult with those in Admissions, Graduate Education, and Online Education to work out viability and logistics. A final proposal would be considered by Academic Affairs and Responsibility, Diversity, Accessibility, and Transparency. Collectively, they would make a recommendation to the Dean who has the final word on behalf of the SDS. Further approvals would be through normal University channels.

Case 2: Private Sector Engagement. A member of Private Sector Relations is presented with the opportunity from a private company to establish a \$20M machine learning initiative that includes a presence on Grounds. This is not a hypothetical opportunity but a real one which the DSI is currently working on. This opportunity is discussed within Private Sector Relations. The company is seeking a research collaboration as well as the opportunity to work with SDS and other students. As such, Career Services is consulted, as is Research Development. Given the ethical considerations, Responsibility, Diversity, Accessibility, and Transparency are consulted

and also sign off on the proposal. A proposal is presented to the Dean who decides to discuss the implications with the Advisory Board (a group of high-level private-sector executives) and the Academic Advisory Board. They endorse the proposal and send it to licensing and ventures for vetting and then to the Provost's office.

Case 3: Research. Research Development identifies a major funding opportunity in biosensors. This is discussed with members of the Center for Data Acquisition, who decide it is worth going forward with an application. As this has an educational component, graduate and undergraduate functional-module team members are consulted. Research Affairs then moves forward with coordinating with other on-Grounds faculty in key departments and beyond and a full collaborative proposal is developed and submitted.

Case 4: Government Relations. The SDS will have a member on the Information Technology Advisory Council (ITAC), a committee of the Virginia Information Technology Agency (VITA). They identify the SDS as a host for a nationwide meeting looking at the data science of cybersecurity in recognition of prior meetings, notably the Governor's Data Analytics organized by the DSI. Events team members consult with other team members in Operational Affairs, and a proposal is prepared, including a budget which is presented first to Strategic Affairs for consideration, additional justification and approval and then to the Dean for approval.

Diversity and Inclusion

Diversity and inclusion are central to our plans for research, education, and general operations.

In the realm of data science, **algorithmic bias** refers to the ways in which AI-informed decisions impact certain demographic, socioeconomic, and geographic groups differently. For example, racial gaps (or overrepresentation) in training data can lead to biased insights from machine learning techniques. **Computing accessibility** refers to the ability of all people to access, interpret and use digital tools regardless of accessibility or visual impairment. In data science, we challenge ourselves to devise interfaces and visualizations that make the power of this emerging discipline available to all. **Data sovereignty** refers to the legal notion that data are governed by the law of the country in which they were encoded. This idea has significant implications for the rights of minority groups and First Nations, as well as for inter-state conflict, war, and cyber security. The ethics course required of all MSDS students covers these topics in detail. We will expand our focus on these and related areas under the auspices of the Center for Ethics and Justice, first through programming and guest lectures, then through hiring, research and grant-seeking.

The DSI team of 19 individuals comprises 10 males and 9 females. One member of our current team is African American. The DSI has a Diversity & Inclusion Council which meets weekly to develop a diversity strategy to govern upcoming hiring. Regardless of SDS formation, the DSI is in a period of rapid hiring, and we are very concerned about being proactive in encouraging members of underrepresented groups to apply for all open positions. It is more than words in a job posting; it is active solicitation of those who might otherwise be intimidated or discouraged from applying. It is devising systems and procedures that guard against implicit bias in

candidate screening. It is a set of diverse committees and hiring officials. It is a welcoming and inclusive feeling upon interview and then as part of the team. We are working hard to assure that the faculty and staff composition of a future School of Data Science reflect diversity in gender, race, ethnicity and other factors. While hiring is the top consideration at the moment, the Diversity & Inclusion Council plans to develop strategies in the areas of admissions and events/programming, among others.

One area in which we have placed particular emphasis is in gender equity in data science. The director has published²³ about best practices. For three years we have connected to the international Women in Data Science initiative to host a Charlottesville a Women in Data Science conference with over 250 attendees²⁴.

A growing number of strategic partnerships and pipeline programs will enhance our capabilities in the area of diversity and inclusion:

- In January we submitted a joint proposal to NSF with **Hampton University** to develop the Virginia Data Corps: linked service learning classes that would develop data science skills using local government and nonprofit organization data sets. Faculty from Spelman College are also involved in this project
- This summer we are supporting students from **Howard University, Spelman College** and the Virginia-North Carolina **Louis Stokes Alliance** to conduct mentored research in Data Science at UVA.
- In August we will launch a student capstone project to investigate housing inequality in partnership with **the Albemarle County Office of Equity & Inclusion**.

We aspire to be a leader in the area of diversity and inclusion both within the field of data science and at the University. Research shows²⁵ that the ethnic diversity of a region impacts its research institution's diversity and that there is evidence that an increase in diversity among faculty at research institutions correlates to increased research expenditures. Programmatic change at the University has the potential to break the cycle of diversity within the institution and the region, thereby attracting a more diverse workforce. We plan to connect with University efforts in this area in order to break this cycle.

V. Students

The school shall have student recruitment and admissions policies and procedures designed to locate and select qualified individuals.

The SDS will build upon the admission procedures, team members, and protocols in use by the DSI. The DSI has an Assistant Director for Academic Affairs who oversees admissions, in

²³ <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002206>

²⁴ <https://datascience.virginia.edu/2019-charlottesville-women-data-science-conference>

²⁵ https://drive.google.com/file/d/15MeH_xD5OEvazObbPeD7Yblhdv3QdS4I/view?usp=sharing

conjunction with the MSDS Program Director. The DSI also recently hired an Admissions and Recruiting Coordinator to build out the Admissions Office. Admissions currently processes approximately 400-500 applications for the MSDS and MSDS dual-degree programs with about a 20-30% admission rate. The DSI Assistant Director for Academic Affairs oversees the review process which includes automated and manual review. All applications are reviewed through automated filtering using multiple data points (data science in action!) before being subjected to human review of all applications. Admissions considers diversity among equally qualified candidates. This is described in an internal DSI document - Admissions Standards and Practices. An expanded version of the process already in place will be used to support the online MSDS admissions process and new degrees as they are approved by the senate and SCHEV.

Student applicants for the residential and online program are recruited through the DSI Website, through graduate program fairs, individual advising hours, diversity events, networking events in various cities, as well as through social media. The DSI also reaches out to UVA undergraduates, encouraging applications to the MSDS program.

Career services is an integral part of the DSI program and done in collaboration with SEAS. We have an Associate Director for Career Connections and Community engagement who is 50% SEAS, 50% DSI and a career placement specialist 100% with the DSI. They coordinate interactions with employers, organize career events, and take students on career treks to high density regions of data science opportunities.

The DSI Team has a relationship of trust with the students brought about in part because of working with the students before, during and after their UVA experience and by being colocated. These features we intend to build into the SDS. Team members and students interact on a daily basis. Initial planning for the SDS building as part of the gift agreement call for a 70,000 sq ft building which will be designed to maximize human interactions. This is described further under resources.

All these operations will be expanded appropriately as part of Academic Affairs in the SDS.

VI. Resources

The school shall have other human, physical, and financial resources sufficient to ensure success in pursuit of its mission.

Physical facilities

The DSI currently occupies the “sheds” Dell-1 and Dell-2 comprising approximately 3,000 sq ft of space, including open space, conference rooms, a few offices and a large display wall. In addition, satellites currently occupy small amounts of space in Fontaine (SOM), the Physical Life Science Building (College), and Ruffner Hall (Curry) . Together, these comprise

approximately 1000 sq ft. The DSI's current budget covers the cost of utilities. As stated above, the gift calls for an SDS building comprising 70,000 sq ft of space. Clearly, growth will occur before a new location is available, and so there are two phases - temporary and permanent. Each is introduced, and we are working with the Office of the Architect for the University on planning and designing future locations.

Temporary

Temporary space of over 4,000 sq ft has been identified on the floor above the Ragged Mountain athletics store on the corner. Plans for configuring that space and a detailed budget are currently being drawn up. Occupancy is expected in late summer 2019. It will also be possible to put a triple-wide trailer next to Dell-2 starting in summer 2020. While not ideal from the point of view of keeping SDS team members and students co-located, it is a workable temporary solution. Funds to support the lease, utilities, etc., on the additional space will be met from funds associated with the gift.

Permanent

Five sites (ordered randomly) are under consideration to provide a 70,000 sq ft building:

- Brandon Avenue next to the new student center
- Emmet/Ivy corridor development
- Current site of Dell-1,2
- West complex
- Stacey Hall site

Details of the building have yet to be worked out, but the DSI team put forward a set of desirable characteristics²⁶ which, it should be noted, coincidentally map closely to the plan put forward to the President by the Committee exploring the Emmet/Ivy corridor site. Whatever the site, the SDS will seek a design to maximize interdisciplinary interaction with other schools, the private sector and various communities. The pro forma budget submitted with this proposal includes operation costs for the building. Bidding for the building architect has begun.

Administration

Since the SDS will operate as a team, the roles and responsibilities of support staff were considered with faculty in section IV and builds upon what is already in place with the DSI. Here we consider only:

Relationship with existing UVA administrative staff responsibilities

The outstanding DSI team will evolve into the expanded team that leads the SDS. How positions will morph with modified responsibilities will be worked out over time and done so in discussions with the Provost's office and HR. This has been the modus operandi as the DSI has expanded. A model that has worked well is considering shared administrative responsibilities across the DSI and other schools and we plan to employ this as much as practical in

²⁶ <https://docs.google.com/document/d/1Q8ojTlx-A-m4EH8mBW5fuifhn6nwehKhX6n5fujcypVg/edit#heading=h.46zoofkbyzr>

relationships between the SDS and other schools. A case in point is career services described above. Joint personnel (via an MOU) work across DSI and SEAS to realize career opportunities for students where there are shared job opportunities. This has worked well and will be continued and expanded into other administrative areas.

Plan for student advisement

Currently the DSI, through the MSDS capstones and presidential fellows, advises students using its own general faculty and faculty from other schools. This model will continue as faculty and students grow. Over time tenured faculty will assume leadership of student advisement roles. With PhD students entering the SDS, they will also play a mentoring role for MSDS and undergraduate students. The details of this arrangement have yet to be worked out.

A plan for undergraduate mentoring, at least for the undergraduate certificate in data science, is being built into the governance structure across all schools. We anticipate there being high demand for these services. SDS will also review being part of the Total Advising²⁷ initiative.

Financial analysis and budget projections

When the idea for the SDS was first presented to the President-elect, one of his concerns was the sustainability of the SDS. To quote, “a large gift, directed at endowments and a building does not make a school sustainable.” At that time the DSI, with enormous help from UVA’s Vice President for Finance, developed a pro forma budget for the SDS (submitted with this document). The budget uses the same template that was developed when the Batten School was formed.

In summary, the five-year projections show the SDS to be sustainable.

Budgetary support

The following list the current and support funds for the SDS:

- Existing DSI endowment - \$13M
- Annual gifts \$0.5 - 1.0M
- Gift - \$120M with matching funds of \$38M as follows:
 - Bicentennial Professorship Endowment Fund \$38M; match \$25M - total \$63M
 - Non-endowed faculty fund to encourage new hires \$3M
 - Bicentennial Fellows Endowment Fund \$26M; match \$13M - total \$39M
 - Visiting Scholars Endowment Fund \$10M
 - Building Fund \$43M
- Current MSDS tuition ~ \$2.5M per year
- Future (after senate/SCHEV approval of new programs)

²⁷ <https://advising.virginia.edu/>

- Return on indirects from research grants - the DSI does not currently receive these, they go to the School where the principal investigator has their appointment
- Tuition from the online MSDS
- Executive education tuition
- Undergraduate tuition
- Joint initiatives with the private sector
- Bootcamps and workshops

Plan for sustainability of the school

Refer to the pro forma budget and the discussion above.

VII. Implementation Plan

The school shall prepare a 5-year implementation plan for faculty recruitment, development of degree programs, student enrollment and growth.

The gift agreement calls for an Implementation Committee (also described above) to be established that meets the requirements associated with the gift. That committee consists of the DSI Director/Acting Dean, UVA President or representative, Provost or representative, Senior VP for Operations or representative, VP for Advancement or representative, donor and donor representatives. The committee will meet quarterly from 2019 through 2021. The first meeting was held on February 7, 2019 and a second meeting is planned for April 5, 2019 as part of the DSI Board meeting. Reports to the Implementation Committee will include an implementation plan for which this document forms a beginning. The implementation plan will be continuously available (a living document) to Deans and other UVA leadership as part of the transparent operation of the SDS.

Timeline

The accompanying spreadsheet²⁸ provides an overview of the timeline for the SDS. These are, of course, best estimates on a very aggressive timeline and depend on Faculty Senate and SCHEV approvals. The competition from other institutions combined with the demand for a trained workforce in data science call for such an accelerated timeline. As the SDS develops and demands become clearer, we will need to be agile and make adjustments to the timeline accordingly.

- The process by which the Dean of the SDS will be appointed is under consideration and will be defined by the President, Provost and Provost-elect. Until then, the DSI Director will continue to coordinate developments in the role of acting Dean.

²⁸ https://docs.google.com/spreadsheets/d/1ycjtQ4LeEsWL-unB-G4edM0gOrF66bgwGCE_-Ahq3QM/edit#gid=0

- The gift calls for 10 endowed chairs at early-, mid-, and late-career stages. The hiring process will begin in earnest, but it will likely be 2-3 years before that full complement is reached. The process to be followed will be similar to what the DSI team is already doing in chairing cluster hires. Search committees comprise DSI faculty and faculty from departments from across Grounds with both a knowledge of the capability of candidates and the ability to determine collaborations that could take place with those appointments.
- SDS needs to hire two persons to be responsible for overseeing the SDS's finance and human resource needs, respectively. These appointments, like all DSI appointments thus far, are expected to be versatile and participate in a variety of functional modules (Fig. 6) as the SDS grows.
- Additional faculty and staff hires will be made on as needed and as funding allows.
- The timeline will be reviewed by the Implementation Committee at each meeting and adjustments made. The timeline will be a public document accessible on the SDS website currently under development.

VIII. Evaluation

The school shall have an explicit process for evaluating and monitoring its overall quality against its mission, goals, and objectives, and for assessing the school's effectiveness in serving its various constituencies.

Evaluation is discussed in the context of the mission (Section I). Of course, we need to use data science as part of that evaluation. The DSI has begun this process, and examples are given in accordance with the major aspects of our mission. All the tools and processes developed within the SDS will be available to other schools should they wish to adopt them.

Train a responsible interdisciplinary workforce

The DSI retains data on our alumni which, over time, describes their career trajectories. We will use these data in concert with a set of evaluation criteria. This will be part of a closed loop system which feeds back into our curricula, teaching methods and areas of research. Those criteria have yet to be established. The process for doing so will be for the Responsibility, Diversity, Accessibility and Transparency, Student Affairs and Research Affairs groups to draft these evaluation criteria which will then be presented to the Advisory Board and the Academic Executive Board (Fig. 6) for consideration. A final list of criteria will be presented to the Dean for consideration and then made broadly known to all SDS stakeholders.

Undertake leading-edge interdisciplinary, open research

Traditional research criteria are not sufficient for the SDS being envisaged. Publications in high impact journals, the success of our PhD students, grant funds are of course important, but do not speak fully to being interdisciplinary and open. Again, these criteria will need to be defined using the same process described above. The DSI is already working with institutions which

have declared themselves open and in collaboration are establishing a set of criteria²⁹ by which to operate. Interdisciplinary criteria will be based on such features as the breadth and scope of research published by SDS faculty, breadth of one's collaboration network (DSI is developing tools so this can be measured) and the types of grants they receive.

Maximize societal benefit

This is the most difficult aspect of the SDS mission to quantitatively evaluate. Again, the process to define evaluation criteria will follow that defined by the other components of the SDS mission. It will include data on usage of research output by others, e.g., data and software reuse, patents filed and licensed, receipt of humanitarian and other types of rewards by the SDS team for services rendered and so on.

Conclusion

The University of Virginia has the opportunity to create a new school in the rapidly emerging field of data science. This rare opportunity takes an already successful Data Science Institute (DSI) and expands its operations to that of a School of Data Science (SDS) providing education for a rapidly expanding student population interested in the field, a research agenda that cuts across all disciplines on Grounds and the opportunity to serve all stakeholders in our university, community, Commonwealth and beyond. We at the DSI feel it is an opportunity to respect the past but address higher education of the future in a way that is sustainable and brings great credit to the University. We hope the reader feels the same.

²⁹ <https://doi.org/10.21955/gatesopenres.1114891.1>

Appendix A The School of Data Science - An Open Scholarly Ecosystem

This appendix introduces the idea of making the proposed School of Data Science an open learning and research environment. What it means to be “open” is defined as well as the motivation and process for doing so. The intent is to raise awareness and seek feedback on the proposal.

Proposed Charge

The proposed School of Data Science (SDS) at the University of Virginia is committed to a culture of ethically responsible openness as its default mode of operation, thereby maximizing inclusion, accessibility and diversity. At the same time, it is recognized that the School must have a business model that is sustainable and provides appropriate security and protection of intellectual property. Pragmatically, being open means operating in a transparent manner and, as far as practical, making the products of the SDS freely and timely available without restriction, provided attribution is given to the producers of those products. Products include, but are not limited to, SDS policies and procedures, educational materials and the complete research lifecycle - lab materials, data, analytics, and the published literature.

Motivation

One of Thomas Jefferson's greatest accomplishments was the founding of the University of Virginia (UVA). A less known achievement was *A Bill for the More General Diffusion of Knowledge*³⁰ presented to the Virginia legislature in 1779. While initially defeated, it laid the groundwork for free public education in the US. To honor that latter achievement, it would seem fitting that UVA continue to lead “the more general diffusion of knowledge.” With the formation of the SDS we have a unique opportunity to do so. We can declare the School of Data Science as an open ecosystem at the point of formation in the same way that the Montreal Neurological Institute (MNI) did³¹, albeit long after the Institute was established. For us, it will be easier as it will be open from the start, not requiring a major cultural shift in an existing organization.

For its part, the federal government is increasingly supporting openness through such resources as open data³², and the funding agencies encourage openness through requirements for open access to the research papers produced, requirement for data sharing plans on grant applications etc. The Human Genome Project was an early open scientific project, and to date it has generated a 140-fold return on investment³³.

³⁰ <https://founders.archives.gov/documents/Jefferson/01-02-02-0132-0004-0079>

³¹ <https://doi.org/10.1126/science.351.6271.329>

³² <https://www.data.gov/>

³³ <https://www.genome.gov/27544383/calculating-the-economic-impact-of-the-human-genome-project/>

Contrary to what one might expect, if MNI is any guide as an early adopter, funding opportunities can increase³⁴, philanthropy increases and more of the best students and faculty want to be part of an open organization - it becomes a movement that stakeholders want to join. The DSI Board has unanimously endorsed the conversion to a School (see letter of support) and declaring that new School an open ecosystem.

Through the proposed School, UVA is very well poised to become a leading US and global academic institution in open scholarship. UVA is already recognized as such through a progressive library and associated initiatives such as the HathiTrust and its Research Center (which provides computational access to four billion pages of text from library collections), the spin-off Center for Open Science³⁵, and the DSI which is operating the Open Data Lab³⁶ and has a faculty handbook which promotes an open culture. UVA could lead a movement that has started to drive the growth and dissemination of knowledge across geographic, temporal, and commercial boundaries. A School of Data Science that declares itself open could be a major economic driver for the Commonwealth of Virginia and beyond while at the same time providing accessible knowledge to a diverse audience.

Impact on Other UVA Schools

From the perspective of other UVA Schools, this initiative provides an opportunity to observe what works and what does not in creating a more open ecosystem - a testbed for possible forthcoming change (or not) at UVA and in higher education. For example, how does it impact enrollment? Such an open ecosystem is not without its political and practical issues. For example, joint appointments³⁷ within Data Science and another School where the criteria for open scholarship are different.

What are the specific issues an open School will address?

- *A culture of open access to knowledge* that is in direct support of the UVA notions of community, discovery and service.
- *Hiring and promotion are traditionally based on bibliometric criteria*, e.g., *H factor*, that do not necessarily represent the maximum value to the education and research enterprise, e.g., shared data, level of collaboration and interdisciplinary research can also be considered.
- *Publication is too slow* - the elapsed time from when manuscripts are submitted to when knowledge is available can be several months to several years slowing the rate of discovery.
- *Access to publications is too expensive* - The cost of providing access to scholarly journals has skyrocketed - journal prices increased 273% over an 18-year period, nearly

³⁴ <http://www.benefunder.com/researchers>, <http://www.orgf.org>.

³⁵ <https://cos.io/>

³⁶ <https://datascience.virginia.edu/pages/expanding-access-data-science>

³⁷ <https://uvapolicy.virginia.edu/policy/PROV-029>

four times faster than inflation at a time when the cost of publishing decreased. Hence knowledge is too expensive and not available to many.

- *Research is suffering from a reproducibility crisis* - too much of the research workflow elements (e.g., data and analytics) associated with research are currently lost. There is no excuse for this in the digital era.
- *Proprietary platforms for data communities are appearing* - The same commercial entities that sell us knowledge at exorbitant prices are beginning to sell us back the data³⁸ we produce upon which that knowledge is based.
- *Teaching materials are costly and inflexible* - High textbook costs hurt students and proprietary content hamstrings teachers. Textbook prices rose over 80% from 2006-2016, three times faster than the rate of inflation³⁹. These costs represent a burden for a significant portion of our students: a third of the latest incoming class qualifies for financial aid and 7% are classified as low-income⁴⁰. 65% of students reported in a nationwide survey that they did not purchase a textbook due to price⁴¹. Proprietary teaching materials discourage innovation in pedagogy, barring customization and adaptation to fit diverse teaching goals and student-centered approaches⁴². Open, living textbooks address these issues and remain current.

How will the School address these issues?

- Maintain a philanthropic fund for promoting openness in the ways described below.
- Openness will be written into the charter, faculty handbook and all other School documents and hence become a founding principle of the School in alignment with the new UVA strategic goals as they emerge.
- All operations of the School will be as transparent as is practical. For example, items such as Board minutes, successful grants, team meetings (suitably redacted) could be made open but the effort is simply not made at present.
- Faculty hiring and promotion criteria will include evidence of open scholarship through collaboration, shared data, shared analytics, prior open publication, community service in open contexts and open teaching materials.
- Student admission will include criteria for evidence of open scholarship and willingness to work in the open.
- The proposed School will provide open infrastructure such as the Open Data Lab and will participate in community initiatives to support open research infrastructures e.g. Wikidata⁴³, Scholia⁴⁴.
- Research will be encouraged to be open throughout the research lifecycle.

³⁸ <http://www.sr.ithaka.org/blog/the-strategic-investments-of-content-providers/>

³⁹ Bureau of Labor Statistics. 20164. Consumer Price Index Databases.

⁴⁰ <http://digital.uvamagazine.org/articles/2020-insight/>

⁴¹ <http://sparcopen.org/wp-content/uploads/2016/01/SPARC-Open-Education-Fact-Sheet.pdf>

⁴² <http://www.slideshare.net/orbitdog1/open-pedagogy-for-elearning-pioneers>

⁴³ https://www.wikidata.org/wiki/Wikidata:Main_Page

⁴⁴ <https://en.wikipedia.org/wiki/Scholia>

- While the research is ongoing, all data, analytics and narrative will be encouraged to be accessible through the resources of the Open Data Lab, the Center for Open Science or other platforms that support research in the open.
 - Once research is completed, UVA's Libra⁴⁵ institutional repository use will be encouraged to preserve all content.
 - All School personnel will have ORCID IDs, and commit to well-designed metadata, and harvesting arrangements to feed tools and services that support open scholarship e.g. Zotero, Protocols.io, Zenodo, Wikidata, VIVO, etc.
- Faculty and students will be encouraged to provide preprints of their articles to recognized preprint servers.
- Publication under open licenses will be expected unless there is strong justification to do otherwise.
- Publication in collaboration with the UVA library using platforms such as Ubiquity or Coko will be encouraged.
- Teaching materials will be as open as possible, while respecting copyright and tuition revenues needed for sustainability.
 - The school will work with the UVA library in support of open e-book services for textbooks and other teaching materials
 - Participate in efforts such as the Provost's Initiative which was recently funded to pilot the creation of an open course platform and seed it with content to be shared within the Commonwealth of Virginia
 - Participate in efforts such as MITRE's Generation AI Nexus, perhaps as an institutional partner on the contribution of selected educational content to the platform
- Work with UVA's Licensing and Ventures Group to maximize openness while retaining intellectual property rights in cases where there is a clear institutional advantage in doing so.
- Encourage private sector participation in open data science at the pre competitive stage where longer term economic value can be seen.
- Continue to encourage and expand UVA's research data management⁴⁶, including to monitor progress, particularly with respect to compliance, use and reuse through the FAIR principles⁴⁷ - Find, Access, Interoperate and Reuse.
- Offer School prizes and possibly Presidential Awards which reward open scholarship.
- Encourage collaboration between and across UVA's schools and collections in an "open by default" manner.
- Encourage public participation in open knowledge projects, e.g., through citizen science and engagement with patient organizations, makerspaces and similar groups.

⁴⁵ <https://www.library.virginia.edu/libra/>

⁴⁶ <https://data.library.virginia.edu/data-management/>

⁴⁷ <https://doi.org/10.1038/sdata.2016.18>

Conclusion

Establishing a new School of Data Science, if endorsed by the university faculty, leadership, BOV and the state, provides a very rare opportunity to start an open knowledge movement from scratch without the burden of an existing culture and associated policies and procedures. We believe it is an opportunity which should not be missed.

Appendix B Support Letters from Deans, Vice President for Research, Internal Collaborators (Departments, Institutes, and Units), and DSI Advisory Board Members



Campbell Hall
P.O. Box 400122
Charlottesville, VA 22904-4122

tel: 434 924 3715
fax: 434 982 2678
arch.virginia.edu

March 28, 2018

Dear President Sullivan

On behalf of the School of Architecture, I am writing to express my enthusiastic support for the proposed new School of Data Science. Data science has had an enormous impact on architecture and the wide range of design disciplines that we are engaged with at the School of Architecture.

Geographic Information Systems (GIS), geospatial analysis, and geodesign are emerging areas of increasing importance at the scales of Landscape Architecture, and Urban Design and Planning, while building information modeling (BIM) and many forms of computational and parametric modeling that we use in the design and analysis of buildings and their systems are increasingly implicated by the information that is embedded in the geometric and spatial models that we are working with. Architecture has witnessed a paradigm shift as the traditional methods of drawing and representation are being displaced by new computational methods and modeling systems—smart three-dimensional models that enable us to do multiple forms of complex analysis and that are also changing the ways in which we are prototyping designs as we directly move from these models to fabrication. Data science is also changing our design process. We now use generative design processes where we start by designing with data. This data operates as a genetic code or a set of changing parameters that we use to hypothesize different outcomes through the manipulation of a structure, building's, landscape's, or city's datasets. It is truly transforming our discipline.

I also believe that as the leading School at UVA in design and visualization, that we have much to contribute to the new School of Data Science by applying our visualization expertise to research in other disciplines, from the data visualization of demographic data in the social sciences, for example, to three dimensional modeling of structures in molecular biology. This new School would enable UVA to position itself at the leading edge of a critically important field that would, in my mind, be certainly beneficial to all of the Schools that comprise the University.

Sincerely,

Dr. Ila Berman
Dean and Edward E. Elson Professor
UVA School of Architecture



**COLLEGE and GRADUATE SCHOOL
of ARTS & SCIENCES**

New Cabell Hall 515
1605 Jefferson Park Avenue | PO Box 400772
Charlottesville, VA 22903

P 434.924.4611
F 434.924.1317
lib4n@virginia.edu

March 21, 2018

Dear President Sullivan,

I write to share my strong support for a new School of Data Science. The field of data science has had—and will continue to have—a vital impact on the fundamentals of a 21st century liberal arts education and on our research enterprise in the College and Graduate School of Arts & Sciences.

As part of the New College Curriculum pilot, our faculty agreed to add a Quantification, Computation, and Data Analysis requirement for all students. This action will better equip students by providing them with the necessary skills to succeed in a rapidly transforming world brimming with data.

Advances in analytics, data architecture, and technology are also highly relevant to our research efforts across the disciplines—including the digital humanities and the social sciences as well as the natural science fields. Partnering with a new School of Data Science with Arts & Sciences leads to exciting possibilities. The capabilities of new instrumentation are immense, producing enormous volumes of data. The field of data science is essential because it allows us to analyze and draw conclusions from data much more quickly and accurately, helping us share our findings for the benefit of society.

I greatly welcome the opportunity of collaborating with a School of Data Science through joint faculty appointments, shared physical spaces, collaborative grants, and common or dual degree programs. This would lift Arts & Sciences and the University as a whole, and establish UVA as a preeminent leader in the space.

Best regards,

A handwritten signature in black ink that reads "Ian Baucom".

Ian Baucom
Buckner W. Clay Dean of Arts & Sciences
University of Virginia



Curry School of Education
417 Emmet Street South
P.O. Box 400260
Charlottesville, VA 22904-4260
Phone: 434.243-5481
Email: pianta@virginia.edu

Robert C. Pianta, Dean

3/31/2018

Teresa Sullivan, Ph.D.
President
University of Virginia

Dear Terry:

As Dean of the Curry School of Education, I write to indicate my support for the formation of a School of Data Science (SDS) at the University of Virginia, which I understand will have a new vision of interdisciplinary leadership and governance. Under this governance structure we will develop a charter for the SDS that includes the following guiding principles:

1. The SDS will provide global leadership in data science through graduate and undergraduate education, professional education, research and translation of research findings into products of benefit to society across the region, Commonwealth, US and internationally.
2. Such leadership will leverage the extraordinary opportunities before us in a way that is unique in organization, physical structure and deliverables and not simply echo what other institutions are doing.
3. As such, the SDS will be truly integrated with other UVA Schools and Centers to create a networked organization. This will be accomplished by, faculty with joint appointments, satellite worker spaces embedded within other schools, collaborative grants and training programs, expanded dual degree programs and joint course development.
4. Particular attention will be paid to 3 as it relates to existing Centers within and across Schools and Pan University Institutes.
5. The SDS will bridge the gap between research findings and development of products by direct relationships with the private sector which may include a physical presence within the SDS to maximize collaboration and expose students directly to career opportunities.
6. A feature of the SDS will be an intramural responsibility to work with central administration to apply data science to recommend improvements to the productivity and wellbeing of UVA and its faculty, staff and students.
7. The SDS will be an open and welcoming environment to a diverse community of stakeholders maintained via scholarships, grants and relationships to a variety of diverse communities locally to worldwide.
8. The SDS will feature a shared governance structure to oversee the development that includes Deans, other university leadership, data science leaders from the academic and private sectors and other members all appointed at the discretion of the UVA President.

I believe there are important opportunities for the Curry School as it collaborates and integrates with the SDS. Leveraging collaboration with SDS and our UVA social science colleagues will enable Curry to leapfrog many of our peers, and establish the fundamental importance of data sciences, analytics, and applications for promoting more effective and equitable p-20 public education.

Best wishes,

A handwritten signature in black ink, appearing to read "Robert Pianta".

Robert Pianta, Ph.D.
Dean, Curry School of Education



DARDEN SCHOOL
of BUSINESS

Scott C. Beardsley
Dean
Charles C. Abbott Professor
of Business Administration

March 22, 2018

Dear President Sullivan,

On behalf of the Darden School of Business, I write to convey my enthusiastic support for a new School of Data Science.

The rapidly burgeoning field of data science is important to our graduates in the Darden School because they would be going into organizations that operate in a world of ever increasing and richer data. These organizations would require their hires to have excellent facility in handling large volumes of data and the basic science of data analytics. As a leading school of graduate business education, data analytics has already become a core skill for our graduates. Indeed, our students are already voting with their feet as the longest waiting lists for our elective courses are in the area of data analytics. Data science would be an imperative for successful management and leadership going forward. The need for more and varied courses and for more rigor and sophistication in our course offerings to our students will only increase in the coming years.

In research and scholarship, the success of our faculty will be affected by their ability to stay abreast of the advances in data architecture, analytics, and machine learning. To this point we have been self-sufficient in meeting our research infrastructure and talent needs in the area of data science. However, with the rapid advances in data science and machine learning, our faculty feel a pressing need to collaborate with experts in other disciplines in order to stay at the forefront of their fields. Further, to be successful in publishing in the leading journals of our business disciplines, our faculty require cutting edge data analytic methods. We have particular interests in large volume data handling, data architecture, applications to business disciplines, and the interface of machine learning and human action.

We look forward to the opportunity for our faculty to participate with faculty in the new school, obtaining unique access to industry and companies, sourcing faculty talent, generating joint research projects, access to shared data, student internship opportunities, and faculty exchange. We anticipate and would welcome PhD and post-doctoral talent and research funding opportunities. Most importantly we hope the new school will inspire new and worthwhile research questions, provide access to interesting data, and enable addressing hitherto intractable problems for our researchers.

The new school in data science would solve the problem of minimum scale that inevitably comes from large volumes of data, which a school like Darden (that is not dedicated to data science) cannot afford or invest in amidst our other priorities. Similarly, a school dedicated to data science could afford to invest in a common research infrastructure that all schools need in data science, but either cannot singly afford or end up duplicating investments across schools. Such an infrastructure would include support for grant writing, cutting edge research equipment, data bases, and a PhD program - all dedicated to data science. There are great efficiencies to be gained from having a school dedicated to data science. A new

school could also allow the university and its various schools to invest in an integrated way in locations beyond Charlottesville – for example Silicon Valley – to enable us to have access to centers of innovation in the areas of machine learning and data science generally. A dedicated school would be a magnet for attracting the very best faculty, research, and student talent in data science to the university and its various schools (including Darden) by demonstrating the necessary commitment and building the infrastructure for excellence in a vital and burgeoning field. Finally, a dedicated school would put the university in a great position for research and educational impact by concentrating efforts, infrastructure, and talent in an integrated way in a strategically vital field.

For all of the above reasons, I believe that a new school of data science is a strategic imperative for the university and I thank you for the opportunity to express our enthusiastic support and to work together in developing the new school.

Sincerely yours,

A handwritten signature in black ink, appearing to read "Scott C. Beardsley".

Scott C. Beardsley
Dean



SCHOOL of ENGINEERING
& APPLIED SCIENCE

31 March 2018

Teresa Sullivan, President
University of Virginia
Madison Hall
Charlottesville, VA

President Sullivan:

On behalf of the School of Engineering, I am pleased to provide my support for a new School of Data Science at the University of Virginia (UVA). Data science is growing at an unprecedented rate and is having a marked impact on our society, both economically and culturally. The most important growth and largest impact are forthcoming – probably no discipline will have greater impact on society over the next 100 years. Having a presence in data science throughout the university is critical for our students and for UVA to contribute to society at the highest level. The proposed School of Data Science will serve this role perfectly at UVA.

I anticipate that the new School of Data Science will be an exceptionally strong partner for the School of Engineering, both in education and research. I foresee hiring joint faculty and developing joint programs between the Schools to provide students with a unique and valuable education, and to provide the knowledge creation through research that will drive the future of our economy.

Please let me know if you have questions or need additional information.

Sincerely,

A handwritten signature in black ink, appearing to read "C. H. Benson".

Craig H. Benson, PhD, PE, NAE
Dean and Hamilton Professor, School of Engineering

Office of the Dean

Thornton Hall | Room A124 | 351 McCormick Road | PO Box 400246 | Charlottesville, VA 22904-4246
P 434.924.3593 | engineering.virginia.edu



P.O. Box 800793
Charlottesville, VA 22908-0793
434-982-4050 | Fax 434-982-0874

Office of the Dean

March 20, 2018

Teresa A. Sullivan, Ph.D.
President, University of Virginia
Madison Hall
P.O. Box 400224
Charlottesville, VA 22904-4224

Dear President Sullivan,

On behalf of the School of Medicine, I write to convey my very strong support for a new School of Data Science.

The emerging field of data science is important to our graduates because of the changing environment in which our research, education, and health care is performed, and the changing job market and daily life regardless of specialty. Taking advantage of – or even being able to navigate – the enormous wealth of health care, genomics and population information requires tools and analysis that was not available a few years ago, and we must be able to perform and teach in that space. Health care analytics promises to become critical for the business and policies of health care.

In our school, research and scholarship will also be affected by advances in data architecture, analytics, and machine learning. We have particular interests in the areas of analytics, bioinformatics, and systems medicine, including the ability to analyze and compare multiple complex data sets with one another to understand mechanisms and potential new therapies and best practices. The rise of “e-health” and “M-health” also provides opportunities and challenges in providing secure data storage and sharing environments for scholarship. We look forward to the opportunity for our faculty to participate with faculty in the new school, including joint appointments, research rotations for doctoral students in DS, internal sabbaticals for faculty to work in DS, and joint educational programs/teaching opportunities such as clinical data sciences, or a MS degree in biomedical data sciences.

Thank you for the opportunity to express our support and to work together in developing the new school.

Sincerely yours,

A handwritten signature in black ink, appearing to read "David S. Wilkes".

David S. Wilkes, MD
Dean, UVA School of Medicine
James Carroll Flippin Professor of Medical Science



Teresa Sullivan, President
University of Virginia
Madison Hall
Charlottesville, VA 22904

31 March 2018

President Sullivan:

On behalf of the Frank Batten School of Leadership and Public Policy, I am pleased to provide my wholehearted support for a new School of Data Science at the University of Virginia. Data science and associated applications in the world of public policy are growing at an unprecedented rate and are having a marked impact on our society, both economically and culturally.

The most important growth and largest impact are yet to come. The foundation for modern policy analysis lies in the ability to parse and work with truly large data. Being able to work with and collaborate with faculty and students in a School of Data Science will be a boon to both the university and our school. The proposed School of Data Science will serve this role perfectly at UVA.

I anticipate that the new School of Data Science will be an exceptionally strong partner for the School of Leadership and Policy, both in education and research. I foresee hiring joint faculty and developing joint programs between the Schools to provide students with a unique and valuable education, and to provide the knowledge creation through research that will shape the policy environment of the future.

Please let me know if you have questions or need additional information.

Sincerely,

A handwritten signature in black ink, appearing to read "Allan C. Stam".

Allan C. Stam
Dean and Professor
Frank Batten School of Leadership and Public Policy
University of Virginia
Charlottesville, VA 22904



April 11, 2018

Carl P. Zeithaml
Dean and F.S. Cornell Professor
of Free Enterprise

President Teresa A. Sullivan
Office of the President
University of Virginia
P.O. Box 400224
Charlottesville, VA 22904

Dear President Sullivan:

Since the founding of the Data Science Institute, McIntire School of Commerce faculty members have been very involved in supporting the research and teaching efforts of the Institute, and a number of the DSI graduate students have been enrolled in classes in the Business Analytics Track of our M.S. in Commerce Program. Our faculty are engaged with the cluster hire efforts supporting data science and analytics programs throughout Grounds, working with several schools to achieve a critical mass of scholars in this general area. These McIntire faculty members are also providing strong leadership in data science research at UVA, with a particular emphasis on scholarship through our Center for Business Analytics. The Center also has a number of major firms as partners in our research and teaching programs.

If the University determines that a School of Data Science is the appropriate path to continue this trajectory and collaboration, then the McIntire School will remain fully engaged in this initiative. We hope that a new school will have a governance structure, collaborative approach, and set of research and teaching programs that truly act as a catalyst to drive the entire University to a leadership position in data analytics among preeminent institutions. Without question, realizing this vision and achieving this goal will require an investment of resources and the recruitment of scholars and data scientists in the schools, including the McIntire School, that currently lead the way in this area at UVA. Analytics is a key competency of the School, both in terms of research and across all of our academic undergraduate and graduate programs, and the opportunity to work with a larger critical mass of colleagues certainly would be beneficial. Our diverse set of courses and programs in business analytics are experiencing tremendous demand and very fast growth. Our primary constraint is the best faculty members to teach in these programs.

Once again, if the University moves in this direction, the McIntire School looks forward to playing an active role in the collaborative development of this broad and very important initiative. We are fully committed to the continued development of our excellence in this area with colleagues from all parts of the University, and we appreciate your efforts to support our efforts.

Sincerely,

A handwritten signature of Carl P. Zeithaml.

Carl P. Zeithaml

CPZ:cbf

CC: Mark Luellen, Vice President for Advancement

University of Virginia, Rouss & Robertson Halls, P.O. Box 400173, Charlottesville, Virginia 22904-4173
434-924-3176 Fax: 434-924-3185 E-mail: czeithaml@virginia.edu
www.commerce.virginia.edu



April 2, 2018

President Teresa Sullivan
P.O. Box 400224
Madison Hall
University of Virginia
Charlottesville, VA 22904-4224

Dear President Sullivan,

As Dean of the School of Nursing at the University of Virginia, I am supportive of the formation of a School of Data Science (SDS), which will have a new vision of interdisciplinary leadership and governance. I join with my fellow deans in the efforts to establish this new School.

I agree with the principles outlined by the SDS thus far and believe that this will be a key driver to more collaboration on a deeper level for UVA. There are innovative opportunities for nursing students and all those interested in health care. In fact, one of our PhD students is completing a master's degree this semester in Data Science and is also winning acclaim for her work with big data related to the microbiome and health. The idea of joint degrees for students and faculty with appointments in several schools will make UVA a leader in big data.

I will work with colleagues to encourage and support this new School and know it will only enhance our ability to discover new knowledge and educate the students of the future.

Sincerely,

A handwritten signature in black ink that reads "Dorrie Fontaine".

Dorrie K. Fontaine, RN, PhD, FAAN
Sadie Heath Cabaniss Professor of Nursing and Dean

UNIVERSITY OF VIRGINIA LIBRARY

February 19, 2019

Dear President Ryan,

The Library supports the development of a School of Data Science at the University of Virginia, because we understand the importance of data science across the disciplines. At the Library, we have been providing support for research data services for some time now, and staff in that area of the Library have been working closely with data science from its inception as an Institute here at UVA. Building on that, we see a new School of Data Science as a key partner in developing 21st-century library services, for example in support of text-mining in different domains. The Library is currently involved in a project to pilot text-mining across the 15 million volumes of the HathiTrust and the journal literature contained in JSTOR and the dark archive Portico—which contains about half of all journal literature in all disciplines. This work involves protected (copyrighted) data and thus dovetails in some significant technical ways with work being done elsewhere at UVA (under Ron Hutchins) on computation with other kinds of sensitive and protected data.

I foresee continued collaboration with Data Science on externally funded grant activities in the Library, joint appointments of library faculty, shared technological infrastructure, and coordinated development of information resources in support of Data Science. In particular, beyond Research Data Services, we see substantial opportunities for collaboration with the Scholars' Lab, which is the Library's digital humanities support unit. Digital Humanities is a place where the University has just launched a new graduate certificate, in which a great deal of student interest has already been evinced: making this one of the areas of focus for Data Science makes a great deal of sense as well.

Finally, we in the Library are very interested in open access scholarship, open data, and open educational resources, as is Phil Bourne. We have been working together on these issues since before Phil arrived here at UVA, since when we have launched Aperio, the Library's open-access journal-publishing infrastructure, and we have collaborated with Data Science to bring Lane Raspberry, resident Wikimedian, to help us identify, describe, and share open data that's being produced here at UVA. The Library will continue to work with Phil and Ron Hutchins and others to complete a census of research data held at UVA, and to promote open sharing of data, teaching materials, scholarship, and research results.

I hope and trust that the University will pursue the very special opportunity presented by this historic gift to found a School of Data Science. Doing so will benefit the Library, the University, the State, and the rest of the world.

Sincerely,



John Unsworth, Dean of Libraries, University Librarian, and Professor of English

ALDERMAN LIBRARY • P.O. BOX 400109 • CHARLOTTESVILLE, VIRGINIA • 22904-4109
PHONE (434) 982-2677 • <http://www.lib.virginia.edu> • FAX (434) 924-1431

October 15, 2018

President James E. Ryan
Madison Hall
1705 University Ave
Charlottesville, VA 22903

Dear President Ryan,

I write to support the creation of a new School of Data Science at the University of Virginia.

We see great demand across the Commonwealth for better analytical skills, from continuing education all the way through to graduate degrees. The School of Continuing and Professional Studies is excited to explore partnerships with the School to help working adults advance in their careers.

We also believe the School's mission of promoting the ethical use of data is an important one and is deeply aligned with the values of the University.

Sincerely,



Alex Hernandez
Dean, School of Continuing and Professional Studies

cc: Philip Bourne



March 30, 2018

Teresa Sullivan
President
University of Virginia
Box 400224
Charlottesville, VA 22904

Dear President Sullivan,

It is with great enthusiasm I provide this letter of support for a new School of Data Science at UVA.

The analysis of large and increasingly complex data sets promises enormous opportunities and insights; however, they will only become fully realized as we develop the tools to manage and analyze them and train a cadre of scholars and practitioners to use them in creative ways. Innovative use of new approaches to datasets will allow entirely novel fields to be created and allow UVA to drive new inquiry and new solutions that impact on society. UVA not only must be a player in this arena - our existing strengths and disciplinary diversity position us to be a leader.

While the prospects are enormous, the need to organize, store, manipulate, and analyze these data sets present major challenges that are limiting progress. Private sector partners are seeking smart investments in the opportunity as are federal sponsors. UVA is already taking on big data and is positioned to become a leader. We have pockets of outstanding research activities that use complex computational approaches, and when brought together they are distinguishing. In addition, the presence of law, business, education, engineering, science, the humanities, medicine, and nursing on one campus allows uncommon synergies that promise to be highly productive in partnership with a new school.

In May 2012, the UVA community came together for a "Big Data Summit." The half-day event was organized and sponsored by the UVA Alliance for Computational Science & Engineering and the Office of the Vice President for Research. Over 170 people attended from 32 different departments. This summit galvanized the community, and its transformative effect is still reverberating across Grounds. It revealed data and analytic parallels and new synergies among diverse disciplines. This ultimately led to the creation of the University's first Pan University Institute: The Data Science Institute.

The institute has brought together the entire computation, analytic, and big data research communities at UVA as well as administrative units like the library and ITS. In addition to the traditional sciences, mathematics, and engineering units, it also includes: the social sciences,

genomics, personalized medicine and medical records, the digital humanities, arts, environmental sciences and sustainability, and translational medicine.

The institute developed curricula, certificates, degrees and majors not only to train quantitative data and information researchers, but also to provide a general, common literacy to anyone whose research touches on data science.

The next exciting stage in the data science evolution at UVA is to develop a school that leverages the transformative scholarship and research opportunities developed over the past 6 years and takes the University in a new exciting direction.

Best regards,



Melur K. (Ram) Ramasubramanian
Vice President for Research
Professor, Mechanical & Aerospace Engineering

cc Mark Luellen

April 3, 2018

Dear President Sullivan and President-Elect Ryan:

On March 23, 2018, the Advisory Board for the University of Virginia (UVA) Data Science Institute (DSI) heard plans for the formation of a School of Data Science. The full board endorses, in the strongest terms, moving forward with the formation of the School as soon as possible.

This is an extraordinary period of opportunity and a time to provide leadership as we move to a data driven economy. As representatives of the private sector engaged in data science across a wide range of industries and sectors, we recognize these opportunities, but also recognize the dangers of not acting. Every day brings new technology, changes in business models, and opportunities to make more data-informed decisions, as well as creating ethical challenges. The time for UVA to set the foundations for becoming a global leader in this area is now—and to wait means giving up a competitive advantage. We are all vested in seeing UVA achieve national and international leadership in data science. The DSI, under the direction of Don Brown and now Phil Bourne, with the help of a very strong team has laid the foundation for achieving this leadership. With the goal of creating a non-traditional, truly interdisciplinary school, we will provide value not only to students of data science, but to all members of the University community by creating an environment that improves data literacy and enables data informed decisions. UVA is well-positioned for this next step.

Studies clearly show the gaps present in the training offered for data driven jobs in Virginia and across the country and the world, but very few institutions are positioned to attack this issue head on with an interdisciplinary approach that fosters creative thinkers with the right quantitative skills. As leaders in this area, we value the DSI approach of project-based learning, interdisciplinary curricula, and team science recognizing the complexity and the depth of a problem. Data science is touching all of us, and UVA needs to provide data literacy for all students and to appropriately train the next generation of practitioners to maximize the number and value to our workforce. This is not only our view, but that of the Business Higher Education (BHEF) and the National Association of Colleges and Employers (NACE), among others. UVA needs robust interdisciplinary research programs where the output of that research is translated into products and services that improve the human condition. Everything we heard regarding the proposed school - guiding principles, organizational structure, educational programs, research initiatives and engagement with the private sector - all lead us to believe that a School of Data Science will position UVA to become that world leader in data science.

We are willing and able to assist in getting the School launched in any way we can.

Sincerely,

Linda Abraham, Co-founder, comScore, Children at UVA

Rob Alexander, CIO, Capital One

Rich Byrne, Senior VP, Programs and Technology, MITRE Corp.

Heidi Lanford, VP, Enterprise Analytics and Data, Red Hat, Class of '91 A&S

Scott Stephenson, President and CEO, Verisk Analytics, Class of '79, Mechanical Engineering

Sam Wholley, Partner, Riviera Partners

Oscar Wood, CEO, NNDATA Class of '95, Electrical Engineering

Appendix C - Other Letters of Support

Email from Alison Booth,
Dear Phil,

It's certainly safe to say that the Scholars' Lab (do note the quaint apostrophe after the s) would welcome collaboration with a School of Data Science. I recommend, as well, that you include language to indicate that there are a range of librarians in various subject areas with technical expertise, some of whom have expertise working with data. So you will probably, with Raf's recommendation, mention Research Data Services, directed Michele Claibourn in the science and engineering library (that's the correct spelling of her name), a political scientist.

Yes, these are ways for Scholars' Lab to collaborate. I give detail designed to discuss these ideas, maybe more than you would want to cut and paste into the statement for the Senate:

- Capstone - e.g., the collective biographies of women project. Note that <http://cbw.iath.virginia.edu> is an IATH project, primarily. Faculty work with both IATH and Scholars' Lab at different stages of their projects. Most faculty in Arts and Sciences won't have \$5K to contract for a capstone project (longstanding humanities faculty essentially don't have research funds). So it would be important to **endow something that people can apply for**. Arts Humanities and Social Science-VP for Research funds are only \$3K, and competitive (two calls per year). I am using my research funds, which I negotiated as director of Scholars' Lab.
- Scholars' Lab works with graduate students as well as faculty in Architecture, Landscape Architecture, and others: large-scale data on archeology and heritage sites. Our two GIS specialists and two AR/VR specialists know the latest on spatial data and cultural heritage informatics.
- Joint seminars and workshops at the intersections of data science and social justice or public humanities, e.g. Resurfacing Black Lives
- PhD student or research fellow shared across SDS and the Scholars' Lab. We can offer semester- or year-long fellowships. Our funding is very limited, and we need to raise funds and endow these fellowships. Arts and Sciences and SDS and the Library jointly should do so!
- Will any SDS students envision academic or teaching positions? SDS could support a digital pedagogy/data fellow who might be mentored in the Scholars' Lab with the Praxis program and other graduate fellows, to prepare them to offer short or long courses in data science and humanities of social sciences (e.g. in media studies; the arts such as sound and music, motion capture, new media installations, etc.; textual analytics; machine learning, AI, and ethics; studies of social media; politics and data journalism)
- Course development grants for faculty in various UVA schools. Scholars' Lab offers technical consultation for [Course Enrichment Grants](#) and the related Research Sprints. It's very possible that faculty will propose courses/projects that intersect closely with Data Science expertise, and we can collaborate in that way.
- Not specific to Scholars' Lab, but to DH@UVA: Developing a deeper bench of faculty able to teach digital humanities will support the Graduate Certificate in DH, which is open to any

graduate student at the University. The Faculty Advisory Committee (chaired by me) for this certificate mentors the students; any Data Science student could participate. At the same time, the resources of SDS would be valuable to Arts & Sciences, Architecture, Education, Nursing, etc. graduate students wishing to take the certificate. This may seem gratuitous for a data science student, but again, some of them may envision academic or library careers.

As you can tell, I would hope a school would regard itself as not just an on-ramp for jobs in the tech industry. And I can tell, by your asking me, that you have real interest in the interconnection with excellent resources of this university.

I'd like to warn against having no tenure in the school. If I'm hearing correctly that that is the plan. It is a dangerous precedent for universities. AAUP has guidelines that UVA officially adheres to.

Best,

Alison

Alison Booth

Professor of English

Academic Director, Scholars' Lab

University of Virginia



SCHOOL of ENGINEERING
& APPLIED SCIENCE

Department of Computer Science

23 February 2019

Rice Hall | 85 Engineer's Way | PO Box 400740 | Charlottesville, VA 22904-4740
P 434.982.2200 | F 434.982.2214 | engineering.virginia.edu

Dear Prof. Bourne:

As you requested, I am writing to convey a Computer Science perspective on the potential of creating a new School of Data Science. These comments largely reiterate my comments at the Feb. 7 Academic Affairs Committee meeting.

Our department has discussed the possibility of a School of Data Science in department faculty meetings, including several with you, so I'm representing a broader, departmental perspective here than just my own as chair. I have also discussed this subject with our department's strategic planning committee.

The phase-1 document outlines the intellectual merit and distinctiveness of the field of Data Science well. It overlaps with many other fields, but it is unique in the way it integrates these to understand data in context and develop methodologies to do this at large scales, and in a way that addresses concerns about provenance throughout the process, from acquisition to drawing conclusions, and in a way that is ethical. The social implications of data are just as important as the analysis. I don't believe these can be holistically addressed without a new unit that is responsible for bringing together these disparate threads.

Many, if not most, academic disciplines emerge as a synthesis of other disciplines, but go on to develop their own unique foundational material. Computer science is an example; it blends electrical engineering and math, and still has strong ties to these areas; AI brings strong ties to statistics. But I think there's a strong case to be made that without CS as a unique discipline, we would have missed out on many innovations, ranging from theoretical to applied. For similar reasons, we support creation of the School of Data Science to help bring this new discipline of Data Science to full fruition, both within UVA and in the world at large. And it is the way data science cuts across so many of our existing schools that I think warrants creation of a school, rather than merely a new department. I really can't imagine how a department would fit into any particular school at UVA.

Creating a School, with the kind of resources enabled by this generous donation, gives UVA an opportunity to establish itself as the international leader in this emerging discipline. This, combined with the incredible resources of UVA as a comprehensive university with great strength in many of the related disciplines, gives students the chance to get a unique education that I don't believe is feasible anywhere else. From what I have seen, UVA is unique in the extent to which it plans to synthesize these disparate issues, from the computational and analytical to the social and ethical, into a coherent body of knowledge. This will make UVA graduates into well-rounded leaders in data science, and from a practical standpoint, create wonderful opportunities for high-impact and fulfilling careers for our students. Altogether, these are exactly the kind of roles that I understand a school should play.

To achieve this vision will require an appropriate balance of fundamental and applied content in both the research and educational programs, an appropriate balance of tenured/tenure-track and general faculty, and most likely faculty with joint appointments between the School of Data Science and other related units, such as CS. It will also require educational programs sized appropriately so that the tuition, in conjunction with endowment revenue, is able to build a critical mass of faculty and scholarship to be a top school internationally, renowned for both scholarship and education, as well as an effective nexus for collaboration within UVA. Because the School of Data Science will depend on collaboration with many other schools, it is important that the School's governance provide a strong voice for faculty from the affected units, to help in selecting a focused research program, hiring appropriate faculty, and building strong educational programs.

In summary, CS is very supportive of a School of Data Science, as long as it is structured to be a world leader in data science research and education and gives an appropriate voice to all stakeholders. We think this is an incredible opportunity for our students and UVA as a whole.

Sincerely,

Digitally signed by Kevin
Skadron 50
Date: 2019.02.23 15:36:51 -05'00'

Kevin Skadron
Harry Douglas Forsyth Professor and Chair of Computer Science
Fellow of the IEEE and the ACM
skadron@virginia.edu, <http://www.cs.virginia.edu/~skadron>



February 22, 2019

Phil Bourne, Ph.D.
Director, UVA Data Science Institute

Dear Phil,

I am happy to write this letter attesting how the Data Science Institute (DSI) has benefited trans-disciplinary research activities at UVA. As you know, I direct the UVA Global Infectious Diseases Institute (GIDI). Like the DSI, we are a Pan-University entity. We were officially launched in July 2017, but began only 18 months prior as a grass-roots initiative among the faculty, following a competitive call for proposals from Provost Tom Katsouleas. We had numerous conversations across Grounds with colleagues and leadership in diverse disciplines including public policy, anthropology, medicine, biophysics, and commerce in order to understand what was needed and what GIDI could do to address it. While we wrestled with these complex issues, DSI gave us insight and support about best practices for building a new, independent research entity at UVA. This letter is my opportunity to thank you specifically for your mentorship. From designing collaborative activities, to creating a leadership structure, to advocating for resources and recognition, you and DSI have been invaluable to GIDI's success. Moreover, the directorship of GIDI and DSI have had a number of conversations about scholarship to mutually benefit our members and, by extension, our Institutes. For instance, last spring we co-sponsored a working lunch about "big data" in Global Infectious Diseases, which led to GIDI members proposing capstone projects for Data Science students. We have followed the DSI model in encouraging scholarship from trainees, who are at the center of many of our Collaborative Seed Grants, and more recently as direct recipients of GIDI's iGrant funding mechanism. Most recently, the four Pan-University Institutes are working together to launch the "Building Bridges" lecture series, which will feature high-profile speakers who work at the interface of society's most pressing issues. The goal of the series is to highlight groundbreaking research that arises from teams that cross traditional disciplinary boundaries, as inspiration and motivation for UVA's transdisciplinary research endeavors.

Going forward, I foresee many opportunities where GIDI might interface with the proposed School of Data Science (SDS). First, GIDI could offer SDS students research opportunities using real-world datasets in global infectious diseases. The capacity to analyze large data sets underlies the most complex problems in global infectious diseases; examples include the spread of a potential pandemic agent via international transport routes, tracking the evolution and transmission of antibiotic resistance in a hospital setting, or modeling the complex metabolic interactions between microbial communities in the GI tract that influence the outcome of an enteric infection. SDS students would have an unprecedented opportunity to work at the interface of two impactful and cutting-edge fields. Second, and relatedly, GIDI and SDS could partner to offer PhD and postdoctoral fellowships at the interface of data science and infectious diseases. As an extension of the DSI Capstones and modeled on the Presidential Fellowships, together we could support co-mentored trainees who will be the next-generation leaders in these fields. Third, GIDI could assist SDS with elements of its educational mission, for instance upper-level coursework in data science and infectious diseases. GIDI and other Pan-University Institutes are not currently allowed to offer degree-granting programs, so partnering with SDS could be a win for trainees who are invested in data science and seek to apply it to global health. Finally, GIDI and SDS could work together to recruit highly talented junior and senior faculty who work at the interface of global infectious diseases and data science. These hires would be transformative to UVA, jump-starting new collaborations with existing faculty,

Global Infectious Diseases Institute
University of Virginia
Box 400301, Charlottesville, VA 22903

 UNIVERSITY OF VIRGINIA
GLOBAL INFECTIOUS DISEASES
INSTITUTE

attracting additional new talent to UVA, and providing new scholarly opportunities to trainees at all levels. Like the other Pan-University Institutes, GIDI cannot hire faculty, but is eager to partner with the Schools to encourage hiring of faculty who will have joint appointments in order to have the greatest impact across Grounds.

The model proposed for the UVA School of Data Science is outward-looking, cross-disciplinary, and inclusive, just like the Pan-University Institutes strive to be. We at GIDI look forward to working with you and your team on this new endeavor.

Sincerely,



Alison Criss, Ph.D.
Associate Professor, Microbiology, Immunology, and Cancer Biology
Director, UVA Global Infectious Diseases Institute

Global Infectious Diseases Institute
University of Virginia
Box 400301, Charlottesville, VA 22903



BRAIN INSTITUTE

Phil Bourne, PhD
Stephenson Chair of Data Science and Director
Data Science Institute
Professor, Department of Biomedical Engineering

Dear Phil,

I am writing to affirm that data science plays a key role in the current function and future plans of the Brain Institute. Our goal is to establish centers of excellence around big neuroscience-related challenges to society: Autism and other neurodevelopmental disorders, Alzheimer's disease and related neurodegenerative disorders, traumatic brain injury and neuro-modulation. Data science will play a key role in each one of these centers.

We are establishing an autism center around two big data projects. Don Brown, the Data Science Institute's founding director and a chaired professor of systems and information engineering in the Department of Engineering Systems and Environment; and Micah Mazurek, associate professor and director of the Supporting Transformative Autism Research (STAR) project, are working with an interdisciplinary team to build a new system for collecting, integrating and analyzing data from multiple sources. They plan to use machine learning and integrated data –to develop personalized therapy approaches for individuals with autism.

In addition, recruitment of Prof. Kevin Pelphrey to UVA has provided access to multimodality (MRI Neuropsychology, genomics) data accumulated serially over a decade that will be treasure trove for ongoing analysis and generation of knowledge. Goal is to identify which children are going to benefit from specific interventions.

In Alzheimer's disease your recruitment of Tim Clarke has provided the glue to developing an Alzheimer's disease Research Center at UVA. Prof. Clarke's goal is to combine data from Prof. Tim Salthouse's Virginia Cognitive Aging Project (VCAP), which is currently one of the largest active longitudinal studies of aging involving comprehensive cognitive assessments in adults ranging from 18 to 99 years of age in the world, with clinical data from Memory disorders clinic at the UVA Health Sciences Center. Over 5,000 adults have participated in Virginia Cognitive Aging Project (VCAP), in three-session (6-8 hour) assessment at least once, with about 2,500 participating at least twice, and about 1300 of them participating three or more times. When these data are combined with clinical data, predictive algorithms for defining risk and outcomes can be developed.

The traumatic brain injury group is also developing a program to create a data matrix of biomechanical forces acting during concussion, neuropsychological tests, balance tests, PET and MRI images collected from subjects, such as special forces members and professional athletes, who are at risk of sustaining repeated sub-concussive and concussive injuries.

1 Box 801330 University of Virginia- Health Sciences Center, Charlottesville, VA 22908-1330, USA

Each of these programs provides opportunities to train undergraduates, graduate students and postdoctoral fellows. In fact the Presidential fellowship program has already supported students such Haoyi Long and Jamie Blair who are using big data and machine learning approaches to study neuroscience.

We foresee an increasing role for data science and data scientists within neuroscience: both in fundamental mechanistic research and in clinical and translational research. It is critical that our students are well trained in data science. A school of data science will be an immense asset to the Brain Institute: in attracting grant funding, faculty, postdoctoral fellows, graduate students. It will greatly increase our ability to compete for resources.

We strongly support formation of a School of Data Science at the University of Virginia.

Sincerely,

 Digitally signed by
Jaideep Kapur 83
Date: 2019.02.18
17:38:04 -05'00'

Jaideep Kapur MBBS, PhD
Eugene Meyer III Professor of Neuroscience
Professor of Neurology
Director, UVA Brain Institute

Hi Phil -

I hope this is within the time frame you need it.

I see great synergies between the School of Data Science and the Environmental Resilience Institute. Both bring together diverse perspectives in teams to address societal issues that no one discipline could do alone. For ERI, tackling major environmental challenges - climate change, water and food security, health impacts - is leveraged by new massive data streams that feed into new modeling and predictive capabilities. These include global remote sensing data, new environmental sensor arrays, complex systems models, visualization tools and participatory games.

ERI has benefited from DSI-funded projects and graduate fellows that fall in this domain. Examples include: using commodity data to understand how global food trade and virtual water transfer affect local water scarcity; refining data-driven complex systems models that underlie participatory games (e.g., Chesapeake Bay Game) and decision-support tools (e.g., Coastal Resilience tool); and using large data streams from new measurement techniques (sensors, unmanned vehicles) that have the potential to revolutionize our understanding of how the earth's surface influences pollution and greenhouse gases. These are just the beginning!

With the new School of Data Science, there is great potential to accelerate our ability to use data and modeling to build on our legacy of interdisciplinary environmental research and secure our place as a global leader in data-driven, solutions-oriented research. A cross-school appointment in Environmental Data would be an excellent opportunity. The successes to date also indicate that there is a tremendous appetite for PhD and post-doctoral fellowships. I could also see an Environmental Data capstone course that would teach the skills needed to integrate large datasets (both spatially and high temporal resolution) into predictive models. These kind of models and visualizations are important for informing decisions on environmental policies and stewardship.

All the best,

Karen

Karen McGlathery
Professor, Department of Environmental Sciences
Director, Environmental Resilience Institute
University of Virginia



SCHOOL *of* ENGINEERING
& APPLIED SCIENCE

Department of Engineering Systems and Environment

Olsson Hall | 151 Engineers Way | PO Box 400747 | Charlottesville, VA 22904-4747

P 434.924.5393 | F 434.982.2972 | engineering.virginia.edu

March 22, 2019

Dear Professor Bourne:

I was pleased to attend the March 21st meeting of the Faculty Senate's Academic Affairs Committee to voice our department's support for the proposed School of Data Science. I would like to expand on my comments in this letter to use as appropriate in the approval process. As stated in the Phase 2 proposal, the Department of Engineering Systems and Environment (ESE - home of UVA's degree programs in Systems Engineering and Civil Engineering) is one of the departments most impacted by the creation of SDS, as one of the departments teaching in the existing Master of Science in Data Science (MSDS) program. We have appreciated the opportunity to discuss SDS and its relation to ESE with you and other leaders of the proposed SDS.

Strength in data science is absolutely essential in advancing knowledge to enable the design, development and operation of engineering systems that support people and communities – the focus of our programs in ESE. These systems rely on the effective use of enormous quantities of data. Given this, our faculty have developed expertise in data science that has allowed us to offer courses and conduct research with the existing Data Science Institute. The formation of the SDS with an interdisciplinary focus and a commitment to work collaboratively with other schools strengthens our programs in ESE. We envision an expansion of joint appointed faculty between ESE and SDS as well as expanded opportunities for ESE and SDS students through a broader, deeper set of course offerings.

In summary, ESE is supportive of a School of Data Science. We stand ready to work with the school to take the university to new levels.

Sincerely,

A handwritten signature in black ink, appearing to read "Brian L. Smith".

Brian L. Smith, Ph.D., P.E., F.ASCE
Professor and Chair