Master of Science
(M.S.) in
Data Science and Analytics

University of Missouri-Columbia

Table of Contents

Executive Summary

This proposal is for a 34-credit, online and mixed mode delivery, multidisciplinary Master of Science degree program in Data Science & Analytics (DSA) at MU. The proposed degree will prepare students with different academic backgrounds to become productive data scientists in Missouri and regionally connected industries. The degree program proposed reflects the diversity of the emerging field of data science by offering emphases in several different related disciplines in advanced data science involving Big Data. It also addresses the MU strategic goals of "implementing innovative curricula that feature MU's interdisciplinary approach to problem solving" and "recruit and retain the best traditional, nontraditional, and distance students." This proposal is the result of a 2014 Mizzou Advantage Grant to the authors.

The program will not require new institutional funding as it is structured to be self-supporting by offering the degree online to a broad audience through Mizzou Online. This degree closely resembles a professional degree, with an estimated *time to degree* of 18 to 24 months and an approximate total cost of $35,000 tuition and fees. The financial structure of the proposal follows the University of Missouri's online degree program funding model for returning fees to the Data Science and Analytics program.

The proposal to offer an M.S. program in data science and analytics is based on the following needs:

- Missouri Employer needs for highly skilled data science professionals
- The State of Missouri's need to maintain a workforce that is competitive for high quality jobs in a global marketplace
- Data Science is an emergent discipline that, by its nature, integrates traditional disciplines. This program is a formal mechanism to leverage prior investments in the computing disciplines across campuses and colleges within each campus.

The proposed program is financially feasible, and will be net-revenue positive by the end of the second year; and completely self-sustaining from year two onward. Market research suggests that data scientists will be the in great demand for the foreseeable future, and the supply of qualified graduates with in-depth technical data science skills is far outpaced by this demand.

The program has already developed a wide set of industry partnerships to form a comprehensive industrial advisory board as witnessed by the support letters included in this proposal. It will continue to leverage these relationships for both student recruitment, as well as allowing students to work on real-world problems from their chosen emphasis area. This aspect of the program will be a differentiator between the MU DSA program and similar data science or analytics programs throughout the country.

Lastly, the program is designed in such a way as to allow "emphasis areas" from the three other UM campuses. Initial conversations with the faculty at all three campuses suggest strong interest and there are plans to begin offering courses from other institutions as soon as the foundational program at MU is up and running. Leveraging the resources from all four campuses is one of the unique advantages of this new Masters program.

1. Introduction

"Transformative uses of data are just around the corner. Precision medicine and other forms of smarter health care delivery, individualized education, and the "Internet of Things" are just a few of the ways in which innovative data science applications will transform our future." – DJ Patil, US Chief Data Scientist[1]

Data science is a rapidly blossoming field of study and career with a highly multidisciplinary characteristic. The confluence of *big data*, massively powerful cloud computing platforms, and need of businesses from all sectors to leverage their data repositories has created a high-growth environment and demand for data scientists. Data scientists routinely leverage tools and techniques from computer science, information systems, advanced statistics, and machine learning. To satisfy the growing need for data scientist who can transform large collections of data into actionable decision making products for their employers, we are proposing the Master of Science in Data Science and Analytics.

This multidisciplinary Data Science and Analytics (DSA) degree program will consist of 34-credit hours of learning in the online and mixed mode format in which students will visit campus one time each academic year for an intensive on site learning experience. The academic program will consist of 19 credit hours of core, fundamental data science courses; followed by 9 credits of emphasis area specific courses and 6 credits of industry relevant case studies and capstone project courses.

Data Science is an emerging discipline that, by its nature, integrates traditional disciplines. The proposed degree program will leverage prior investments in the computing disciplines across campuses and colleges within each campus. The MU Informatics Institute will coordinate this collaborative degree program by leveraging existing courses from Computer Science, Journalism, and Information Science & Learning Technologies Departments to deliver the various core and emphasis area course. Existing courses will be adapted to the online format, and new courses that are properly focused and structured for the DSA program will be developed.

"Medical researchers are now building analytic tools that turn raw data into actionable intelligence. Some are conducting vast surveys of vital statistics; others are mining existing databases of medical treatments and outcomes. They're looking for patterns, and the hope to find the best treatments for your specific case of your specific ailment." -- Special Report: "Hacking the Human OS: How Big Data Will Transform Medicine and Health", IEEE Spectrum, June 2015

[1] https://www.whitehouse.gov/blog/2015/02/19/memo-american-people-us-chief-data-scientist-dr-dj-patil

Data Science and Analytics represents the single greatest opportunity to develop a highly skilled workforce in a generation and to gain a long-term economic advantage for the state. To assure the State of Missouri can capture its fair share of this expertise and economic opportunity, we plan to offer a Masters degree program in Data Science & Analytics (DSA).

The proposed M.S. program in data science and analytics will address these needs:

- Missouri employers' needs for highly skilled data science professionals
- The State's need to maintain a workforce that can compete for high quality jobs in a global marketplace.

The DSA program will be guided by an Executive Director, an Industrial Advisory Board, a University of Missouri Advisory Board, and two operational co-directors. The Director of Academic Program for DSA will coordinate the course content and delivery standards among the various academic departments and campuses to ensure a consistent, high-quality, rigorous degree program. The Director for Outreach and Industry Relations will focus on external development and collaboration with industry and business partners.

The two Co-Directors will work closely together to assure the program is financially sound and to continually improve and enhance the program. They will ensure current and relevant technologies are readily available for course instructors to integrate into learning modules so that students receive training on the latest tools, techniques, and technology. The necessary program startup costs have already been secured by the Executive Director, and the program will be financially self-sustaining after the first year (see Section 3.B for full financial models, including costs and expenses).

2. Fit with University Mission and Other Academic Programs

2.A. Alignment with Mission and Goals

The proposed degree program will contribute toward a number of goals at the department, college and campus levels. The proposed program aligns with these from MU's Strategic Operating Plan available at: http://strategicoperatingplan.missouri.edu

1. Strengthen interdisciplinary and experiential learning for Mizzou's undergraduate, graduate, and professional students:
   1.1. Grow undergraduate and graduate collaborative and interdisciplinary research and creative activity with faculty by supporting a data-heavy community that allows all faculty and students the chance to see the benefits of Big Data research, education, and implementation.
   1.2. Implement innovative curricula that feature MU's interdisciplinary approach to problem solving. The proposed program will use a "modularized" course delivery model to leverage existing courseware from various disciplines and add substantial Big Data analytics components to each course for the state-of-the-art technology used in the industry.

2. Recruit and retain the best traditional, non-traditional and distance students, as well as the best faculty and staff. The DSA program will provide data analytics training opportunities for graduate students in traditional degree programs to increase their marketability in both academic and industrial sectors. This will attract students who will select MU's traditional programs over others and promote MU's strategic goals:
   2.1. Encourage and reward effective interdisciplinary work by faculty and staff by providing adequate faculty compensation and creating an administrative infrastructure to support several disciplines and areas.
   2.2. Optimize faculty impact in teaching, research, outreach, and economic development by providing atmospheres of collaboration between faculty, students, business, and industry; by creating teaching opportunities for current PhD students; and by increasing the University's AAU profile through the integration of cutting edge research, online education, and transformative community outreach.
3. Ensure that MU's revenue model allows for strategic investments and leverages MU's strengths to drive state and regional economic development. The DSA Masters is in a unique position to meet state, national, and international market demand by offering an online and mixed-mode delivery method. The program allows for additional emphasis areas to be added in the future, which will increase net revenues, as well as provide data analytics services to the private sector.

We propose to implement innovative curricula that feature MU's interdisciplinary approach to problem solving. The proposed program will use a "modularized" course delivery model to leverage existing courseware from various disciplines and add substantial Big Data analytics components to engage state-of-the-art technology used in the marketplace.

The proposed program will recruit and retain the best traditional, non-traditional and distance students. We will provide data analytics training opportunities for graduate students in traditional degree programs to increase their marketability in both academic and industrial sectors. This will attract students who will select MU's traditional programs over others because of this data analytics training opportunity.

The proposed DSA Masters will help to recruit, develop, and retain faculty and staff to promote MU's strategic goals. First, we will encourage and reward effective interdisciplinary work by faculty and staff by providing adequate faculty compensation and creating an administrative infrastructure to support several disciplines and areas. Second, we will optimize faculty impact in teaching, research, outreach, and economic development by providing atmospheres of collaboration between faculty, students, business, and industry. We will create teaching opportunities for current PhD students, and increase the University's AAU profile through the integration of cutting edge research, online education, and transformative community outreach.

Sustainability is essential to ensure that MU's revenue model allows for strategic investments and leverages MU's strengths to drive state and regional economic development. The DSA is in a unique position to meet state, national, and international

market demand by offering an online and mixed-mode delivery method. The program allows for additional emphasis areas to be added in the future, which will increase net revenues, as well as provide data analytics services to the private sector. Those that might not be able to afford a full time data scientist may still benefit from the skills of MU faculty, staff, and students. As the program grows from its inception at MU, toward integration with other campuses, there will be emphasis areas centered at UMSL, UMKC and S&T. We describe these partnerships in more detail later in the proposal.

2.B. Duplication and Collaboration Within Campus and Across System

Even though the media has concentrated on Big Data as it pertains to business, virtually all parts of the economy are seeing the benefit of using Big Data for new knowledge and discovery. *MU is in a unique position to leverage its research strengths across many disciplines, including health care, veterinary medicine, life sciences, journalism, geography, natural resources, education, and computer science.* Students will have the opportunity to receive advanced training in data science that is transferrable across disciplines; AND study in depth in their particular emphasis area. By building upon existing expertise in these fields, MU can be a leader in data science training.

Compared with other programs around the US, the proposed DSA Masters at the University of Missouri features emphasis areas that are specific to industry vertical specializations. The opportunity to apply data science in the context of a particular set of domain specific problems takes advantage of the strengths across the University of Missouri.

Currently, no public institution in the state of Missouri offers a Masters in Data Science and Analytics. There is a business focused Masters program at St. Louis University, but this program provides an education at a level of abstraction "up" from computing. Its target student population has less technical knowledge than students targeted by the proposed DSA Masters. Similar, Missouri University of Science and Technology is part of a distance and continuing education program centered in business and information technology.

The S&T program could, in fact, offer an emphasis area in the DSA Masters at some point, and we have initiated these conversations with folks operating that program. The S&T Certificate Program is the closest offering to the proposed DSA Masters, but S&T lacks the vertical domain integration and computing education breadth of this proposed, integrated Masters program. It is also true that a certificate program is less comprehensive than a Masters program. Truman State University offers an online graduate certificate in Data Science as part of their distance learning initiative. Missouri State University is in the initial stages of creating an emphasis area for business analytics in the College of Business for Management undergraduates.

There are a few programs that have similar organizational structures to the proposed DSA Masters, but which, upon careful review of the curriculum, do not compete in the same market space at all. These include the Executive MBA Program at MU, the HMI Masters program at MU and the Masters in computer science offered by MU. Potentially

duplicate programs, which we have evaluated as serving distinct, different markets, are presented in the table below.

Institution	Program	Degree	On-line	Emphasis
St. Louis University	School for Professional Studies	M.S. in Applied Analytics	No	Business School Focus
Missouri S&T	Distance and Continuing Education	Graduate Certificate	Yes	Business School and Computing Focus
Truman State University	Distance and Continuing Education	Graduate Certificate	Yes	Computation & Statistics
Missouri State University	[planned] College of Business	Emphasis Area in Undergraduate Management	No	Business School

Collaboration

The DSA program plans to recruit professionals from different industries, as noted in our letters of support and advisory board. MU Campus collaborators include the School of Journalism, the School of Information Science and Learning Technologies, the College of Engineering, and the MU Informatics Institute. We will also be working with other UM System Campuses, including UMSL, S&T and UMKC. This program will train students to handle peta-scale raw data sets and convert them into actionable decisions using the state-of-the-art Big Data technologies with industry-driven problems. It is fundamentally different from traditional computer science programs, which mainly focus on algorithm design and software development. It is also different from informatics programs that are domain driven in either biology or health care without data-intensive training. The student populations are significantly different between the proposed DSA program and existing programs.

Although this proposal is focused on the MU campus to ensure timely offering of the degree program, we have been working with the three sister campuses in the University of Missouri System to allow maximum leveraging of existing faculty, expertise, and resources to eventually achieve a UM-System collaborative degree. The logistics in offering a system-wide degree program would require substantial efforts that would delay the initial offering of the degree program. Based on the inputs from our sister campuses, the structure of the proposed degree has been designed to enable the other campuses to participate – likely by offering additional emphasis areas and providing modules for the core courses.

Potential emphasis areas include healthcare from UMKC, business analytics from Missouri S&T and UMSL. Since October 2015, four campuses have been collaborating in developing intercampus courses and sharing curriculum in the area of Big Data analytics fostered by the UM System. Once the DSA program is launched in Fall 2016 from MU campus, we will start working on new emphasis areas with our sister campuses based on the year-long collaboration in intercampus course development and sharing.

3. Business-Related Criteria and Justification

3.A. Market Analysis

3.A.1. Need for Program

"Advances in information technologies are transforming the fabric of our society and data represents a transformative new currency for science, engineering, education and commerce." "The work we do today will lay the groundwork for new enterprises, promote economic growth, improve our citizens' quality of life, and fortify the foundations for U.S. competitiveness for decades to come." - Farnam Jahanian, NSF "Big Data" Congressional Briefing, May 2012

Companies determine where to locate new businesses and business expansions only after considering the availability of an appropriately skilled workforce. Companies like Cerner, ExpressScripts, MasterCard, and IBM have located all or part of their firms' big data efforts in Missouri because of our strong, foundationally skilled workforce. In many respects, Missouri workers are highly skilled, lower cost alternatives to workforces on the east or west coasts of the USA. However, the "big data" skills and knowledge that are in high demand in the workforce are lacking, and Missouri needs to build its own workforce because we are not a high amenity state, and therefore cannot expect to draw substantial numbers of highly skilled data science workers from other regions. A Masters in Data Science and Analytics is an essential step to maximize entrepreneurship and corporate relocation and fill the demand for highly skilled data scientists in Missouri's future.

Big data, machine learning, and predictive data analytics have been hailed as the fourth paradigm of science - allowing researchers to extract insights from both scientific instruments and computational simulations. Machine learning has yielded new insights into health risks and the spread of disease via analysis of social networks, internet search queries, and hospital data. It is also vital in areas as diverse as high-energy physics and molecular biology. The US Bureau of Labor Statistics (bls.gov) predicts that over 50,000 new, highly skilled data science positions will be created over the next decade in the USA[2]. However, at present Missouri is projected to garner less than 100 of these jobs[3] annually. Our current labor force is surely a significant factor. More of these great jobs in the future of data science are surely Missouri's for the taking, if we can build up our own workforce.

As with successive generations of other large-scale scientific instruments, each new generation of advanced computing brings new capabilities, along with technical design challenges and economic trade-offs. Broadly speaking, data-generation capabilities in most science domains are growing more rapidly than compute capabilities, causing these domains to become data-intensive. High-performance computers and big-data systems are tied inextricably to the broader computing ecosystem and its designs and markets. They also support national-security needs and economic competitiveness in ways that distinguish them from most other scientific instruments."

---

2 http://www.bls.gov/careeroutlook/2013/fall/art01.pdf

3 https://www.missourieconomy.org/

If Missouri is to maintain and grow its significance in this new economy the proposed Masters degree in Data Science and Analytics is an essential component in Missouri's workforce development strategy. The proposed Masters degree in Data Science & Analytics is designed to provide students the necessary knowledge and skill sets across various disciplines needed to meet the high industry demand for data scientists. All students will take "Core Courses" that provide a foundational learning and introduce state-of-the-art technology in Big Data, database design, data ethics, and visualization of high-dimensional and high-volume data.

A 2011 study published by the McKinsey Global Institute indicated that by 2018 the US could face a shortage of as many as 190,000 workers with "deep analytical skills". The study also predicted a work-force gap of 1.5 million managers and analysts with the skills to decipher and translate data patterns for decision-making. Leading state universities in states that have previously succeeded in growing new jobs in the technology sector are already taking action to meet these labor market demands. In January, 2014, the School of Information at the University of California at Berkeley started what is billed as the first all-online Masters degree in information and data science. In Fall of 2014, the University of Indiana began offering a Masters in Data Science to residential students. Our mixed-mode delivery model captures the value of a University of Missouri campus experience, and identity cohorts while maintaining the flexibility required of working professionals.

Many industries are experiencing two momentous shifts: first to cloud-based cyberinfrastructure and second to big data driven business analytics. The convergence of these trends in nearly every industry that relies on computational systems for competitive advantage has solidified the need for a new generation of data scientists who are positioned to exploit tomorrow's computational infrastructure for business intelligence. This requires a unique foundational skill-set from a variety of disciplines including informatics, computing and information sciences, statistics, and more. The Data Science & Analytics program will position the University of Missouri as a leader in the training and development of data scientists in the region, the nation, and potentially internationally based on the mixed-mode course delivery and program structure. The Data Science & Analytics program will offer Missouri companies a local resource of highly trained data scientists, thereby advancing their competitive position and thus providing economic benefits to the state.

3.A.2 Benefits to the State of Missouri

University System Benefits

The very nature of data cuts across academic disciplines, allowing for multiple collaborations and expanded potential for knowledge discovery. The proposed degree structure allows for future growth - emphasis areas to be added to meet the expanding analytic needs of not only the university, but also business, industry, and government. As the program develops, other campuses of the University of Missouri are expected to participate and offer degrees through this program. The proposed degree will help create

a community of faculty from Engineering, Journalism, Education, and other disciplines dedicated to big data and analytics and interested in graduate level Data Science and Analytics education. At present there is no locus for interaction related to graduate level DSA education within the University of Missouri System.

Research, Teaching and University System Benefits

The proposed DSA Masters has a number of positive benefits for the University of Missouri system's business model. For example, this program will enable the University to hire highly skilled data scientists for university oriented business development. Bringing Data Science to each campus in the UM System would provide a significant competitive advantage for Missouri business and industry, state agencies and universities and thus enhance the state's economic development.

To achieve these benefits, we will develop a clearinghouse (online or otherwise) for matching research projects with DSA Masters students' Case Study and Capstone studies. The aim is to bring Big Data Analytics to campus wide activities. Specifically, these types of capacities will include:

- Data Recon (What data do we have? What is the quality of the data? What's the fit between the data a scientist has and the questions a scientist has?)
- Data Janitorial (Cleaning and staging the data for analysis)
- Data Structuring (Reshaping the data to support analysis, exploration and visualization)
- Visualization (Presenting the data to customers, stakeholders and the public)

The key opportunity is building a capacity for people to work with the messy, unstructured data and produce useful analytics based on this type of data. This has benefits to Industry that have an interest in processing and making sense of data from social media, discussion forums or other online environments. Essentially, we would be building a data intensive computing resource focused on messy data problems.

Educating Missouri's Students

To understand real world Big Data issues in context, students will select three courses within an emphasis area. This offers in-depth analyses and training on data analytic techniques, issues, and problems students will face within a given concentration area. Students will take a Case Study course to gain hands-on experience with large data sets and use the latest technology and techniques. A Capstone project will then enable them to refine and demonstrate the knowledge and skills they learned throughout the program. Both courses will provide students with mentoring from faculty, as well as insight from industry partners.

We propose to offer a 34-credit, online and mixed mode (courses with one campus visit per academic year) delivery, multidisciplinary Masters degree program in Data Science & Analytics (DSA). The proposed degree will prepare students with different backgrounds to become productive data scientists in Missouri and regionally connected industries. The degree program proposed reflects the diversity of the emerging field of data science by offering emphases in several different disciplines in advanced data science involving

Big Data. This degree closely resembles a professional degree, with an estimated *time to degree* of 18 to 24 months with an approximate total cost of $35,000 tuition and fees. The financial structure of the proposal follows the University of Missouri's online degree program funding model for returning fees to the Data Science and Analytics program.

Letters of Support

We have received numerous letters of support from industry partners and colleagues, who indicate that they would leverage the DSA Masters for their own firm's workforce development. Consistently, our partners express gratitude for the endeavor we have undertaken to increase the availability of a technically skilled data science workforce. They acknowledge the huge and surging demand for data scientist in the age of big data; where data is transforming so many industries. Our industry partners from consulting entities, such as KPMG, Deloitte Consulting, and Silicon Valley Data Science have highlighted the trend of increasing importance of data science for there customers to maintain competitive advantage.

Letter Writer	**Role**	**Industry Affiliation**
Harlen D. Hays	Sr. Manager, Quantitative Research and Biostatistics	Cerner
Rounsley, Steve	Genomics Lead, Seeds and Traits R&D	Dow Agrosciences
Glen Schuster	SVP and Chief Technology Officer	Centene
Heather Nelson	Solution Architect	Silicon Valley Data Science
Paul Pancoast		Deloitte Consulting
Chase Davis	Deputy Editor, Interactive News	New York Times
John Lee	Director, Data & Analytics	KPMG
Valinda Kennedy	IBM Chicago Cloud Academic, Entrepreneur and Developer Gladiator	IBM
Tom Henry	Chief Data Officer & VP	Express Scripts
Don Etling	Senior Vice President/Senior Partner	FleishmanHillard

"We are living in an era of tectonic shifts in the way information is conveyed and received. … I feel it is vital for the University of Missouri – whose cornerstone is the world's preeminent school of journalism – to be fully engaged in this mission-critical discipline."

Don Etling
Senior Vice President/Senior Partner
FleishmanHillard, Inc.

"... we have seen first-hand the march that organizations [are making] from data naive, to have analytics drive their decision making."

John H Lee, PhD
Director, Data and Analytics
KPMG LLP

We receive similar sentiment from our industry partners who are focused in particular business sectors such as biomedical, journalism, and computing. Each of our partners sees the emergence of data science as the next wave of transformation for their industry, similar to the disruptive nature of the Internet on businesses in the previous decades.

"Through collaborations between Cerner Corporation, the University of Missouri, and the Tiger Institute we have identified the criticality of turning data into action. As this need continues to grow, it will require a diverse workforce with a specialization in Data Science to help disrupt the global Health Care crisis. This workforce will be developed through specialized programs such as the proposed Masters in Data Science and Analytics degree at the University of Missouri ... "

Harlan Hays
Sr. Manager, Quantitative Research and Biostatistics
Cerner

3.A.3. Student Demand for Program

Previously outlined labor market analysis, technology industry analysis and the strong letters of support from industry advisory board members demonstrate there is a market need for the proposed program. Specific projections are presented in tables 1a, 1b and 1c.

Table 1a. Student Enrollment Projections (anticipated total number of students enrolled in program during the fall semester of given year).

Year	1	2	3	4	5
Full-Time	40	90	100	110	140
Part-Time	0	0	0	0	0
Total	40	90	100	110	140

Table 1b. Student Enrollment Projections (anticipated number of students enrolled during the fall semester of given year who were new to campus).

Year	1	2	3	4	5
Full-Time	40	50	50	60	80
Part-Time	0	0	0	0	0
Total	40	50	50	60	80

Table 1c. Projected Number of Degrees Awarded

Year	1	2	3	4	5	6	7	8	9	10
# of Degrees Awarded	0	40	50	50	60	80	100	100	100	100

3.B. Financial Projections

3.B.1. Additional Resources Needed

Our budget outlines a need for a website, some space renovations and computing equipment for the co-directors. There is also a budget for staff, supplies and marketing. These costs are justified as usual and customary costs for operating a degree program at MU. Our budget projects that these costs will be well covered by program revenue.

3.B.2. Revenue

Revenue needs will be met through base tuition, and program differential fee outlined in the financial projections. The per-credit hour fees include 354.25 in tuition and a 682.11 differential fee. Our total revenue year one will be $663,267.84, based on the enrollment projections in table 1a; increasing to $2,374,108.95 in year five.

3.B.3. Net Revenue[4]

In year two, we project $736,518 in net revenue; by the end of year five, the annual net revenue will be approximately 1.2 million dollars.

In addition to the $663,267.84 in projected first year revenue, the Executive Director of the proposed Masters degree in Data Science and Analytics has secured both internal and external funding to start-up of the program. Internal support included a 2014-2015 Mizzou Advantage grant for $80,000 to develop Core Courses. In 2014, Dr. Shyu was also awarded $600,000 from the National Science Foundation to acquire high performance supercomputing equipment. This is only a part of the computing and network research cyberinfrastructure that will support Big Data analytics, research, and teaching activities. These investments have helped pave a solid foundation for the startup of the DSA program. It also conveys the commitment to, and appreciation of the need for, the proposed degree program, and the willingness of federal funding agencies to partner in helping increase the data science knowledge and skills abilities of Missouri's workforce. Next, we outline the costs the are deducted from gross revenue to arrive at net revenue.

One-time start-up funds will be required in the first year to gradually cover faculty funding and space requirements, including online delivery facilities, office renovations initial administrative costs. We have allocated $146,000 for Year 1. These funds include $111,000 in space renovations and $15,000 for an initial Website design and deployment. Additionally, funds for equipment will cover the cost of office computers and equipment necessary for conducting business operations.

[4] For more details, please see the DSA Masters Budget Worksheet

Recurring expenses are associated with the day-to-day operation of the program, instructional costs, cloud computing infrastructure for student projects, website and marketing, and staff positions. The current faculty and GTA costs are tied directly to the number of students projected into the program, and will vary with enrollments. Staff and Director positions are set to be relatively stable over the startup period to ensure continuity in the development of a world-class Masters in Data Science and Analytics.

Personnel cost includes (1) Administration (Executive Director, Co-Directors), (2) Supporting Staff (academic advisor, fiscal staff, and cyber engineer), (3) Teaching Faculty, and (4) Graduate Teaching Assistants (GTA).

Administration: The Executive Director is currently supported by the MU Informatics Institute. Therefore, no budget is requested. The Co-Directors will be funded 50% of salaries, including benefits, to support their efforts to lead the operations of the program. $162,444 is based on current compensations of Drs. Sean Goggins and Grant Scott.

Supporting Staff: To support student activities, an Academic Advisor ($55,000 annual salary) is budgeted at 0.5 FTE in Year 1 and 1.0 FTE Year 2 and beyond. The budget includes $20,000 for a half-time (0.5 FTE) fiscal officer. A full-time $40,000 position will be half-funded by another interdisciplinary program, such as MU Informatics Institute. To provide proper cyberinfrastructure for Big Data analytics, a 0.5 FTE cyber engineer is budgeted to provide cloud environments, both commercial (Amazon.com, Google, IBM, etc.) and in-house (campus computing) with a $70,000 annual salary ($35,000 budgeted). To accommodate the expected enrollment and program growth, the cyber engineer position will transition from 0.5 FTE to 1.0 FTE ($80,000) in year 3 and beyond.

Teaching Faculty: To ensure delivery of core courses, the DSA program has allocated funding for six instructors ($10,000 per course plus benefits) per year (total of around $80,000 per year). These instructors will work with the Co-Directors and emphasis area Associate Directors to ensure that their courses meet the Big Data and analytics rigor the program requires.

GTA: Graduate students associated with the program may be supported by graduate teaching assistant funds. We will recruit top-notch doctoral students from emphasis area programs (Informatics, Computer Science/Engineering, Journalism, and Information Science) to serve as Graduate Teaching Assistants (GTAs). One quarter-time (0.25 FTE) GTA will be assigned to a course for every 25 students. The approximate monthly salary is $1000 for a 0.25 FTE GTA. This rate reflects the need of including student insurance and other fees. The GTA budget in Year 1 is $54,000 as budgeted and in the following years $126,000 is the annual cost for GTAs. These collaborative efforts among participating academic units and the DSA program will enhance MU's AAU metrics.

Cloud Computing: Cyberinfrastructure for course learning and projects will be partially subsidized through partnerships and programs with key industry partners. Using MU's alumni network, we have worked with major commercial cloud providers, such as

Amazon.com and Google, to achieve discounted rates and free resources for the students. We have budgeted $1000 per student annually to support the computing needs, and to help fund the depreciation/maintenance costs for our local high performance computing clusters.

Operating: $20,000 annual budget is allocated to support basic operations for the DSA main office including copies, meeting materials, week-long residential activities, etc. We have budgeted $75,000 for marketing in each of the first five years.

Professional Development: Due to the nature of fast-evolving technologies in Big Data analytics, it is essential to allocate funds ($15,000/year) to ensure that relevant faculty, staff, and GTAs have the necessary ongoing training in emerging data science techniques and technologies. These funds may help pay for Big Data and analytics conferences, workshops, and materials; and help bring outside experts to campus for more cost-effective onsite training.

Web Design and Promotion Materials: The DSA program differs from traditional degree programs in terms of recruitment. Because the prospective student is 'online' we will rely heavily on web-based promotion and electronic media to attract students and industry partners. Budget for these materials fall under the $75,000 marketing budget.

Mizzou Online Overhead: Because of the program's online delivery mode of instruction, all courses will be administered through Mizzou Online. The Mizzou Online tuition rate is equal to that of an in-state graduate student, or $354.25 per credit hour ($1062.75 per three-credit course). To offset the additional expenses involved in offering a highly technical and resource heavy program, students will also be charged a $682.11 per credit hour differential fee. To compensate for online delivery services, Mizzou Online will receive the standard 40% return on the Mizzou Online tuition rate ($425.10 per 3-credit course, which breaks down as 25% to Mizzou Online directly, and 15% through Mizzou Online to the Provost), while the program will receive 60% of the Mizzou Online tuition rate and 100% of the differential fee. This combined tuition and differential fee is competitive with other online data science programs.

3.B.4. Financial and Academic Viability

Table 3.b.a. Financial Projections for Proposed Program for Years 1 Through 5.

	Year 1	Year 2	Year 3	Year 4	Year 5
1. Expenses per year					
A. One-time					
New/Renovated Space	$111,000				
Equipment	$20,000				
Initial Web Design/Build	$15,000				
Total one-time	**$146,000**	**$0**	**$0**	**$0**	**$0**
B. Recurring					
Faculty	$81,222	$82,846	$85,128	$87,467	$89,866
Co-Directors	$162,444	$165,692	$170,256	$174,933	$179,732
Staff	$74,454	$75,943	$137,472	$141,249	$145,124
GTA	$54,000	$123,930	$137,700	$156,060	$174,420
Equipment	$0	$0	$0	$0	$0
Operating	$20,000	$20,000	$20,000	$20,000	$20,000
Mizzou Online Overhead	$90,687	$219,349	$250,186	$281,955	$324,606
Professional Development	$15,000	$15,000	$15,000	$15,000	$15,000
Marketing	$75,000	$75,000	$75,000	$75,000	$75,000
Computing	$40,000	$90,000	$100,000	$112,000	$140,000
Total expenses (A+B)	**$758,806**	**$867,761**	**$990,741**	**$1,063,665**	**$1,163,747**
2. Revenue per year[5]					
Tuition & Differential Fee	**$663,268**	**$1,604,279**	**$1,829,811**	**$2,062,170**	**$2,374,109**
3. Net revenue	**-$95,539**	**$736,519**	**$839,070**	**$998,505**	**$1,210,362**
4. Cumulative Revenue	**-$95,539**	**$640,980**	**$1,480,050**	**$2,478,555**	**$3,688,917**

Table 3.b.b. Enrollment at the End of Year 5 for the Program to Be Financially and Academically Viable.

Enrollment Status	Full-Time	Part-Time	Total
Number of Students	**50**	**0**	**50**

[5] The DSA Masters Budget Worksheet, in Appendix C includes a detailed breakdown for these revenue projections. Most simplistically, they are (number of students) X (Number of Credits) X (Tuition + Differential Fee).

3.C. Business and Marketing Plan: Recruiting and Retaining Students

Program marketing will focus on regional and national data science centers. Regionally, we will focus data science markets in the state of Missouri (KC, St. Louis and Springfield), Chicagoland, Oklahoma (Tulsa), Arkansas (Wal-Mart), Tennessee (Nashville) and Kentucky (Louisville and Cincinnati metro). National markets we will focus on are Washington D.C., New York City, San Francisco, Minneapolis, Dallas, Phoenix and Atlanta.

Our marketing will be delivered through three main channels:

1. Internet Search

- Google Adwords will be purchased, driving focused candidate students to our site
- A Website for the program will be developed and include search engine optimization to make our program be returned in the first few items for searches on "Data Science" and "Data Science Education".

2. Vendor displays at key national / international industrial and international conferences

- Example: Techweek conferences in Chicago and Kansas City, and other cities
    - Techweek is a conference series that moves through six cities each year, showcasing entrepreneurship, innovation, and emerging technologies.
    - Big data related companies and technologies are a constant theme at Techweek conferences
- Various industry leading cyberinfrastructure providers hold conferences, such as Amazon's AWS:reInvent. These conferences are network with working professionals within the larger data science and analytics community.

3. Alumni publications reach a wide variety of Mizzou alumni.

Many colleges have specialized alumni publications, as well as the general alumni publications. These publications represent a very specific audience with already existing ties to the Mizzou community, and in some cases the college/department responsible whom are delivering the emphasis area DSA courses.

Table 1a, in Section 3.A.2 shows our initial five-year enrollment projections. The DSA Masters program is designed to be a technically rigorous degree, allowing our graduates to continue evolve as data scientists throughout their careers. Word-of-mouth among graduates, and their management, will build the reputation of the program; thereby creating a culture of advocates in the industrial segment.

As the program matures, we will maintain a program website with update news and events, as well as an electronic alumni and industry partner newsletter. We will continue to establish a vendor presence at relevant conferences that have demonstrated good applicant leads.

We have budgeted $75,000 for marketing in the first 5 years to cover the cost of the marketing activities as described above. The marketing budget will be allocated at the discretion of the Co-Directors. The budget in Section 3.B, includes this allocation, as well as a steady-state marketing effort in the out years for continual electronic publications and participation in select Big Data related conferences.

The DSA Program is designed as an Executive-style degree for working professionals with its mix-mode delivery of mostly online classes and yearly campus visits. We will be leveraging modern online collaboration tools to ensure students stay engaged with the instructors and their cohort students (classmates).

We are structuring the program as follows to ensure the program is optimally aligned to the typical working professional participating in online courses.

- 8 week courses
    - During a standard UM semester, an online course will start on the 1st, 5th, and 9th week of the semester
    - = 3 courses per semester
    - The working students can focus on at most two active courses at any one time.
- Courses organized into learning modules that foster a learning community among the cohort class.

We will work with our industry and academic emphasis area partners to build a world-class, rigorous program, which is self-sustaining. The Director of Outreach and Industry Relations will work with our industry partners to find excellent students within their organizations, whose DSA tuition cost may be fully or partially subsidized by the company. Additionally, through continual professional networking among the program Directors, Industry partners, and program alumni; we will leverage a network of advocates to recruit students to the program. The various associated DSA faculty, including core courses and emphasis area faculty, will leverage their networks to advocate the program and attract qualified students. Our continuous marketing efforts and the reputation for producing quality graduates will ensure we achieve and consistently maintain our enrollment goals.

4. Institutional Capacity

Financial:

The DSA Program will be financially self-sustaining, as can be seen in the Section 3.B Budget. The potential student pool for the program is not limited by campus physical presence, as the program is an online degree for working professionals. This drastically increases the number of potential students, as they may participate in the course from anywhere in the Missouri, the country, or even the world. The program costs each student pays will adequately support the expenditures of the program for said student.

Faculty:

There is faculty available to teach the core courses in the first year, and emphasis areas have committed faculty resources for years two and beyond. Longer term, we aim to target

cluster hires supported by Mizzou Advantage or MUSOP, for both core courses and the emphasis areas.

5. Program Characteristics

5.A. Learning Outcomes

Graduates will be able to individually acquire and stage large data sets, design and conduct experiments, and analyze results for complex data analytical problems using their foundational and specialized data science tools and techniques; taking a problem from conceptualization stage through to the production of data-derived business intelligence.

The special skills the graduating students will acquire or possess include:

- Real-world experience in applying state-of-the-art data science tools and techniques to solve industry, academic, and/or business data and decision-making challenges.
- A clear understanding of the ethics and security mechanisms required to safeguard large-scale data collections that contain sensitive and critical information.
- A comprehensive understanding of modern data analytics, statistical analysis, and visualization tools that facilitate timely, large data analysis.
- A solid foundational understanding of database systems, database design, and information retrieval; allowing exploitation of a broad spectrum of data repositories and streaming data systems.
- A demonstrated ability to effectively communicate to a broad audience the relevant information derived from large data collections using a variety of visualization and presentation methods. Students will be able to convey the meanings behind specific data analysis techniques to audiences of various technical knowledge.
- Training in the latest data analytic methods and tools; including fundamental and advanced statistical and mathematical principles upon which advanced data analysis techniques are built (machine learning, pattern recognition, data mining, etc.).
- Specialized, advanced training in a chosen Emphasis Area, such as biotechnology, high-performance computing, strategic communications, human centered data science design, etc.

5.A.1 Emphasis Area Learning Objectives

In this proposal learning objectives are included for all four emphasis areas. Additional emphasis areas may be developed at a later time.

1. High Performance Computing Emphasis Learning Objectives

Graduates of the Master of Science in Data Science and Analytics who pursue the *High Performance Computing* (HPC) emphasis area will achieve the following educational objectives, in addition to the core program objectives while becoming immersed in Big Data computational ecosystems.

- Students will have an in depth understanding of the state-of-the-art technologies which enable big data analytics and high performance computing; such that they can successfully investigate the data and analytical needs, then guide the decision making process on deployments into HPC infrastructure.
- Students will acquire knowledge to exploit cloud-based computing infrastructure, including virtualization, distributed architectures, on-demand resource scaling, container technology, and other cloud-based computing concepts in support of Big Data management, processing, and analytics.
- Students will have a thorough understanding of advanced technologies and techniques in Big Data analytics which facilitate the extraction of new data intelligence using state-of-the-art, leading analytical platforms.
- Students will gain a solid understanding of techniques for exploiting advanced co-processing hardware, including *graphics processing units* (GPU) and *many-core units* (e.g., Intel Phi) to achieve cost effective, massively parallel data analytics

2. Human Centered Data Science Design

Students will develop a deep understanding of the theoretical foundations and hands-on experience necessary to understand the strengths and limitations of different analytical methods.

- Human Centered Data Science Design combines both the technical (databases, social networking, data mining, and text mining) and social (economic, ethical, policy, and political) aspects of data analytics.
- Students will build an understanding of the complex interplay between the decisions made during the collection, curation, and transformation steps in the information lifecycle, and their impact on the analytical methods that should be employed.

3. Biotechnology

In addition to the core program objectives, graduates of the Masters of Science in Data Science and Analytics who pursue the *Biotechnology* emphasis area will achieve the following educational objectives:

- Possess an in depth understanding of the data analytics needs in biotechnology industry and healthcare systems in the US and worldwide;
- Have the ability to practice their analytic skill sets to applications in agriculture, human medicine, and health information systems in genomics, proteomics, phonemics, and electronic medical records;
- Be able to present and interpret their data analytic results with actionable plans in biotech and healthcare industry.

4. Strategic Communication and Data Journalism

Graduates of the Master of Science in Data Science and Analytics who pursue the *Strategic Communication and Data Journalism* emphasis area will achieve the following educational objectives, in addition to the core program objectives.

- Students will have in-depth capabilities and understanding of big data management, including gathering and interpreting customer and viewer behavior patterns and developing strategies to enhance marketing objectives for organizations and clients.
- Students will have highly-marketable skills in analyzing media markets and will be able to develop sophisticated methodologies to optimize usage of apps, social networks, and other technologies for media businesses and brands.
- Students will be able to apply their analytic skills to understanding and optimizing patterns of search, programmatic advertising buying and behavioral targeting.
- Students will have a deep understanding of issues of privacy and ethics in obtaining and utilizing data from a broad range of sources.
- Students will be able to obtain and analyze publicly-available data in a variety of structured and unstructured formats. As such, they will develop an understanding of open-records laws and how to effectively use them.
- Students will develop the skills necessary to work as a data journalist for a news organization.
- As data journalism and data science /analytics are constantly evolving, students will develop strategies that enable them to continue to learn on the job.
- Students will develop an understanding of their audiences and how to best communicate with them.

5.B. Program Structure

We propose to offer a 34-credit, online and mixed mode (courses with one campus visit per academic year) delivery, multidisciplinary Masters degree program in Data Science & Analytics (DSA). The proposed degree will prepare students with different backgrounds to become productive data scientists in Missouri and regionally connected industries. The degree program proposed reflects the diversity of the emerging field of data science by offering emphases in several different disciplines in advanced data science involving Big Data. This degree closely resembles a professional degree, with an estimated *time to degree* of 24 months with an approximate total cost of $35,000 tuition and fees. The financial structure of the proposal follows the University of Missouri's online degree program funding model for returning fees to the Data Science and Analytics program.

The proposed Masters degree in Data Science & Analytics is designed to provide students the necessary knowledge and skill sets across various disciplines needed to meet the high industry demand for data scientists. All students will take "Core Courses" that will provide a foundation of knowledge and an introduction to state-of-the-art technology in Big Data, database design, data ethics, and visualization of high-dimensional and high-volume data.

To understand real-world Big Data issues in context, students will select three courses in an emphasis area. These elective courses will support in-depth analyses and training on data analytic techniques, issues, and problems students will face within a given Emphasis

area. Students will take a Case Study course to gain hands-on experience with large data sets and use the relevant technology and techniques. A Capstone project will enable students to refine and demonstrate knowledge and skills learned throughout the program. Both courses will provide students with mentoring from faculty, as well as insight from industry partners.

Total credits required for graduation: 34

Residency requirements, if any: None

General education
Total credits for general education courses: 19

Courses (specific course or distribution area and credit hours):

Degree Program Course Offerings	
Year 1	
Fall Semester	Spring Semester
Introduction to Data Analytics (3)	Data Mining & Information Retrieval (3)
Database and Analytics (3)	Big Data Security (3)
Data and Information Ethics (1)	Data Visualization (3)
Summer Semester	
Statistical & Mathematical Foundations of Data Analytics (3)	Case Studies (3)
Year 2	
Fall Semester	Spring Semester
Emphasis Area Course 1	Emphasis Area Course 3
Emphasis Area Course 2	Capstone

Emphasis Area requirements
Total credits specific to degree: 9

Courses (specific course or distribution area and credit hours):

Biotechnology Emphasis	
Course	Hrs
INFOINST 8005 *Bioinformatics Tools*	3
INFOINST 8700 *High-throughput Biomedical Data Analysis*	3
INFOINST 8720 *Biomedical Information Mining and Interpretation*	3

High Performance Computing Emphasis	
Course	Hrs
CMP_SC 7001 *Cloud Computing*	3
CMP_SC 7001 *Big Data Analysis*	3
CMP_SC 8850 *GPU and Multi-Core Computing*	3

Strategic Communications & Data Journalism Emphasis	
Course	Hrs
JOURN 7263 *Interacting Advertising II*	3
JOURN 7236 *Psychology of Advertising*	3
JOURN 7248 *Media Strategy Planning*	3
JOURN 7430 *Computer-Assisted Reporting*	3
JOURN 7430 *Advanced Data Journalism*	3
JOURN 7440 *Mapping for Stories and Graphics*	3

Human Centered Data Science Design Emphasis	
Course	Hrs
IS_LT 9410 *Introduction to Human Centered Data Science Design*	3
IS_LT 7310 *Seminar in Human-Computer Interaction*	3
IS_LT 9410 *Learning Analytics*	3
IS_LT 9410 *Digital Humanities and Information*	3
IS_LT 9410 *Metadata*	3
IS_LT 9410 Learning Analytics	3

Free elective credits

Total free elective credits: 0. All non-core credits are specified by the emphasis areas.

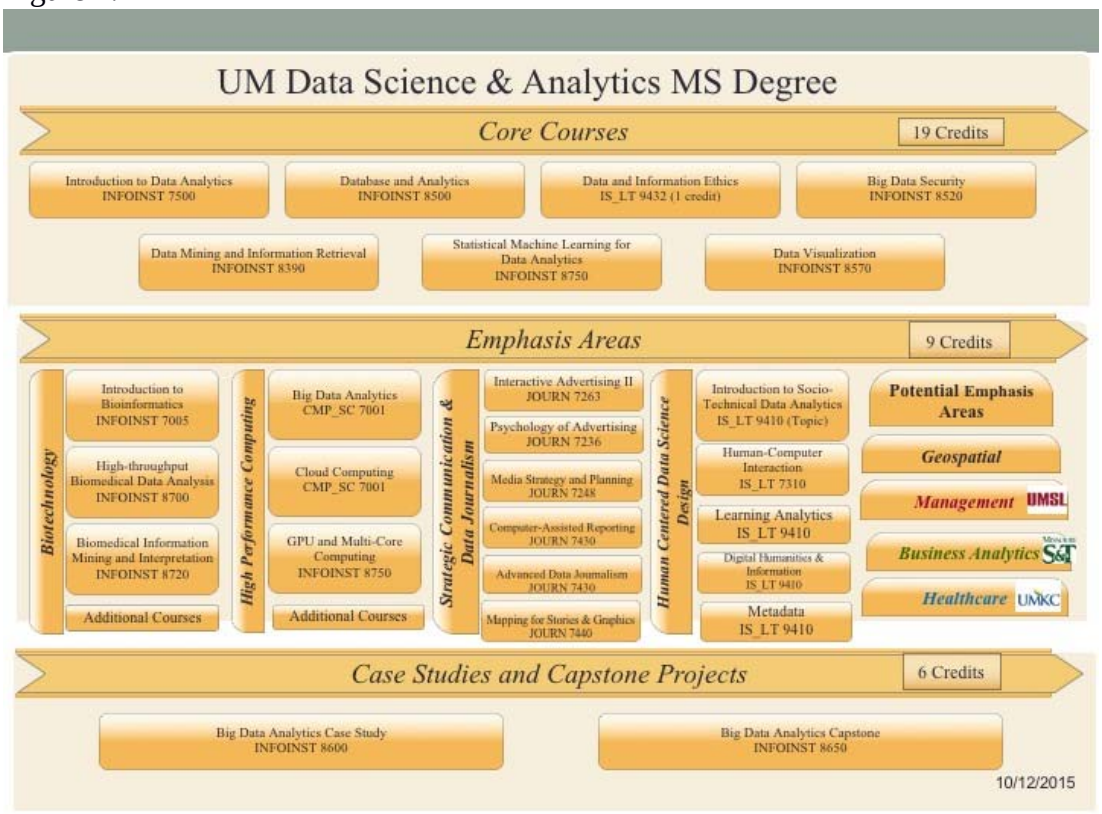
Requirement for thesis, internship or other capstone experience: A three-credit capstone project is required for graduation.

5.C. Program Design and Content

The curriculum evolved over the course of a two-year program design period. The Executive Director and Co-Directors each met with faculty in the core computing curriculum areas several times during that two-year period, while outlining and conceptualizing how the curriculum would be mated with program outcomes. The incorporation of emphasis area program directors, a program advisory board, and a curriculum committee will ensure the ongoing alignment of new and revised curriculum offerings with program outcomes. Industry advisory board members will ensure that the program outcomes remain relevant to industry over time.

The curriculum is described in figure 1, showing the specific sequence that courses in the program will be offered in. The DSA Masters is divided into three primary curricular components: Core curriculum, Emphasis Areas and Capstone Project. The sequence was developed with a few fundamental principles of interdisciplinary education in mind. First, we need to create a solid base of common knowledge for students with different backgrounds. The core curriculum serves this purpose. Knowing that the students we recruit will have diverse experiences and prior education, the first semester of core curriculum is less technically focused. The second semester is more technically focused. Second, the emphasis areas then each build on the core skills, helping students focus their data science knowledge in a particular industry or cross industry specialization. Finally, the capstone project is an independently executed data science project that combines core discipline knowledge with emphasis area knowledge to solve a real industry challenge.

Figure 1.



5.D. Program Goals and Assessment

Program goals are outlined throughout the other parts of this proposal. Learning outcomes will be assessed for each course using a combination of pre and post tests, student surveys and peer reviews of curriculum. Mizzou Online has existing templates and processes in place for these assessments, and we will build on those. Our aim is 100% retention, and we expect over 100 graduates annually in years 3 through 10. Employment projections also suggest that we will experience nearly 100% placement for graduates of the program.

5.E. Student Preparation

The program is designed as an *Executive-style* graduate program for working professionals. We expect that students will be currently working in a data science or analytics related position. The typical DSA student matriculated into the program will be seeking to broaden their exposure to the state-of-the-art tools and techniques of data science, as well as gain deep technical proficiency applying data science and analytics to their chosen emphasis area. We will also accommodate recently graduated students, who are directly transitioning from an undergraduate degree in a related field (such as Computer Science, Statistics, etc.). Students seeking admission to the program and currently working in an industry data-driven job will be expected to have two years of job experience in data science or analytics. Traditional graduate students, will be required to meet a minimum GRE score for admission into the program in lieu of industry experience. Transcripts and letters of reference and support from employers will also be considerations for admission.

Given the inherent multidisciplinary nature of data science and analytics, we expect students to have a diverse background. As such, not all students will have the necessary fundamental knowledge and skills necessary for a rigorous and thorough program. To accommodate this diversity, which we consider healthy for the livelihood of the program, we will provide self-paced learning resources covering key fundamental tools and techniques. These learning resources will be in the form of *Boot-camp* courses for:

- R, the industry-standard, open-source statistical programming language
- Python, a widely used procedural and object oriented programming language
- SQL, the industry-standard language to query relational database management systems
- Linear Algebra, the fundamental mathematical principles upon which much of the advanced data science and analytical technical concepts are developed

Once students are matriculated into the program, these resources will be made available, and students will be expected to pass *pre-test* during their initial *Introduction to Data Analytics* course.

5.F. Faculty and Administration

MU has strong expertise in computing and informatics, disciplines closely related to Data Science. The expertise is spread across a number of colleges and departments and is best organized through the proposed multidisciplinary structure. Relevant research expertise at MU covers the whole spectrum of computing and informatics and extends into a wide range of academic programs.

The DSA program will be lead by two Co-Directors - one for Curriculum and the other for Industry Relations.

The Co-Director of Curriculum will focus on internal development and coordination of the program among the collaborating academic units at MU as well as the intra-campus partnerships. The Co-Director for Curriculum is responsible for coordinating the course topics and learning objectives, as well as ensuring the emphasis areas are high-quality and rigorous.

The Co-Director for Industry Relations will focus on external development and collaboration with industry and business partners. These activities will focus on three important aspects of the DSA program; 1) the Industrial Advisory Board, 2) the integration of real-world problems and projects into case-studies and capstone courses, and 3) effective partnerships with leading big-data cyberinfrastructure companies to ensure student access to state-of-the-art technologies for course work.

The two Co-Directors will work closely together to assure the program is financially sound and to continually improve and enhance the program. They will ensure current and relevant technologies are readily available for course instructors to integrate into learning modules so that to students receive training on the latest tools, techniques, and technology.

Additionally, an Executive Board for the program will be composed of the Executive Director, the Co-Directors, and one Associate Director for each emphasis area. The Associate Directors will represent the emphasis area faculty and offering department/college. The coordination of the Associate Directors and the core program directors will ensure that the core courses continue to provide the necessary fundamentals which enable the optimal delivery of the emphasis area courses.

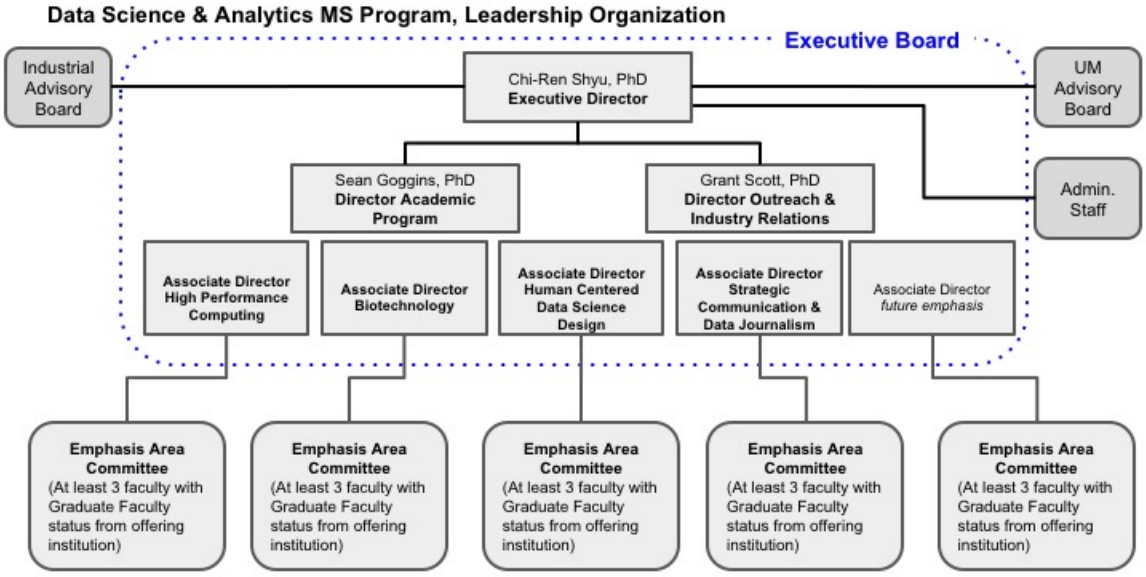
Industry Advisory Board (IAB)
A board comprised of industry and business experts in companies vital to Data Science and Analytics has been formed to help assure the program is relevant and responsive to the marketplace, and help ensure the student's learning experience is relevant to "real-world" technologies, problems and challenges.

Industry Advisory Board		
Member	**Title**	**Company**
Harlen D. Hays	Sr. Manager, Quantitative Research and Biostatistics	Cerner
Steve Rounsley	Genomics Lead, Seeds and Traits R&D	Dow
Tom Henry	Chief Data Officer	Express Scripts
Glen Schuster	SVP and Chief Technology Officer	Centene
Paul Pancoast	Physician and Clinician Consulting and Analytics	Deloitte Consulting
Jordan Tigani	Staff Software Engineer	Google
Chase Davis	Deputy Editor, Interactive News	New York Times
Heather Nelson	Solution Architect	Silicon Valley Data Science
John Lee	Director, Data & Analytics	KPMG
John Everson	Director Information Security & Compliance	DISH Network
Don Etling	Senior VP/ Senior Partner	FleischmanHillard
Valinda Kennedy	IBM Cloud Academic, Entrepreneur and Developer Gladiator	IBM
Ray Brown		FairCom
Janet Roberts	Data Journalist	Thomson Reuters

The value of collaboration and the criticality of the degree program was noted in one of the industry representative's response, when invited to serve on the IAB:

> "Through collaborations between Cerner Corporation, the University of Missouri, and the Tiger Institute we have identified the criticality of turning data into action. As this need continues to grow, it will require a diverse workforce with a specialization in Data Science to help disrupt the global Health Care crisis. This workforce will be developed through specialized programs such as the proposed Masters in Data Science and Analytics degree at the University of Missouri.

University of Missouri Advisory Board (UMAB)
A board comprised of University of Missouri System members has been formed to facilitate collaboration within the four-campus University of Missouri System. This board will be composed of members from each campus, and various schools/departments on each campus, to provide DSA program directors with input regarding the needs of the campuses and the cooperative delivery of the multi-campus degree program.



Name and position of the individual who will be responsible for the success of this program
Dr. Chi-Ren Shyu, Shumaker Endowed Professor; Chairman, Electrical and Computer Engineering; Director, MU Informatics Institute

5.G. Alumni and Employer Survey

(1) Alumni surveys will be conducted annually using a web-based assessment and evaluation method to determine alumni satisfaction and the scope of employment positions and industries. A combination of qualitative and quantitative responses will be used for gathering the information.

(2) A web-based assessment and evaluation method will be used to collect information from the employers of the graduates to determine the value of the DSA program.

5.H. Program Accreditation
At present, there is no accreditation of graduate programs in Data Science and Analytics.

Appendices

Appendix A: Detailed Course Descriptions

Course Description for New Courses

Abbreviated course descriptions for the required data science and analytics courses in the program curriculum are as follows:

Core Courses

INFOINST 7500 - Introduction to Data Analytics (to be developed)
First course in Data Science and Analytics MS degree and graduate certificate programs. The objective of the course is to give students a broad overview of the various aspects of data analytics such as accessing, cleansing, modeling, visualizing, and interpreting data. Students will have hands on training in Python, R, SAS, SPSS, and other open-source analytic tools. Concurrent technologies in Big Data ecosystem will be introduced with applications in the emphasis areas.

INFOINST/STAT 87550 – Statistical and Mathematical Foundation for Data Analytics (to be developed)
An introductory statistics class designed to build the mathematical foundation for students dealing with Big Data phenomena. Topics include reviews of probability, data sampling, data summarization, sampling distributions, statistical inference, statistical pattern analysis, hypothesis testing, regression, and nonparametric inference. Students will have Big Data projects using R, SAS, or SPSS for their course projects using Hadoop.

INFOINST 8500 - Database & Analytics (to be adapted for program; moved online)
Covers the Fundamental concepts of current database systems and query methods with emphasis on relational model and non-relational techniques in Big Data environments. Topics include entity-relationship model, relational algebra, indexing, query optimization, normal forms, tuning, security, NoSQL, and data analytics skills in both relational and non-relational environments. Project work involves modern relational DBMS systems and NoSQL environments.

IS_LT 9423 - Data and Information Ethics (already offered as IS_LT 9423: Information Ethics, 3 cr. hr. mixed-mode delivery. IS_LT 9423 Information Ethics will become "IS_LT Data and Information Ethics." The course will be changed to a modular format containing a series of units targeted to the different "Emphasis Areas." Students will select from the various modules to create a 1 cr. hr. package suited to their Emphasis Area) Introduces the ethics related to Big Data in industry, business, academia, and research settings. Students will learn the social, ethical, legal and policy issues that underpin the big data phenomenon. Discussions and case studies will help guard against the repetition of known mistakes and inadequate preparation. The course content will follow the guidelines to be developed by the Council for Big Data, Ethics, and Society.

INFOINST 8570 – Data Visualization
Data visualization broadly covers transforming multidimensional and time varying datasets to dynamic visual representations and encodings that facilitate exploratory data mining, knowledge discovery, improved understanding, summarization, structural modeling, collaboration and decision making using interactive methods.

INFOINST 8520 – Big Data Security (to be developed)
This course provides and overview state-of-the-art topics in Big Data Security, looking at data collection (smartphones, sensors, the Web), data storage and processing (scalable relational databases, Hadoop, Spark, etc.), extracting structured data from unstructured data, systems issues (exploiting multicore, security). Securing sensitive data, personal data and behavioral data while ensuring a respect for privacy will be a focus point in the course.

CMP_SC 8390 - Information Retrieval (Plan to offer online format)
Theory and techniques for the modeling, indexing, and retrieval of text-based and multimedia databases. Topics include introduction to different information retrieval models, retrieval evaluation, query languages, query operations, and indexing/searching methods.

Advanced Core Courses
INFOINST 8600 - Big Data Analytics Case Study (to be developed)
Using a case-study approach, students will engage in discussions on a variety of big data topics relevant to their emphasis area and the realm of Big Data. This course will help students generate ideas and prepare them for the Big Data Capstone

INFOINST 8650 - Big Data Capstone (to be developed)
As a culmination of their data science experience, students will engage in a Big Data Capstone project, incorporating the theoretical knowledge of coursework learned throughout the program into a hands-on practical data project. Students will work closely with faculty mentors and leaders in business, industry, and government to solve real-world Big Data issues in an applied setting.

Emphasis Areas

Biotechnology Area Courses
INFOINST 7005 - Introduction to Bioinformatics (Plan to move to online format)
Will provide the introduction to the current directions in bioinformatics through the computational tools available to research community. Students will learn how to efficiently apply the tools and software packages as well as analyze and visualize the results.

INFOINST 8700 – High-throughput Biomedical Data Analysis (Course to be developed)
To enable students to analyze and understand their own biomedical data, this course will include: large scale microarray data analysis, normalization, hierarchical clustering, k-means clustering, bi-clustering, GO analysis and Gene Set Enrichment Analysis. The

course will cover tools and techniques for read mapping, analysis of RNA-Seq and ChIP-Seq data, assembly of short reads into contigs, SNP analysis, as well as emerging file formats.

INFOINST 8720 – Biomedical Information Mining and Interpretation (to be developed)
Students will learn the principles of data mining techniques for use in biomedical data analysis, especially in genotype and phenotype analysis. Topics include selection of data sources, identification of proper data mining algorithms, locating cloud computing sources using industry standards (IBM BlueMix, Amazon Cloud, etc.), and evaluation and interpretation of the results with industry-driven tasks.

High Performance Computing Emphasis Area Course
CMP_SC 7001 (Topic) - Cloud Computing (Adapt to online delivery)
This course will explore the principles that integrate computing theories and information technologies with the design, programming and application of distributed systems. The course topics will familiarize students with distributed system models and enabling technologies; virtual machines and virtualization of clusters, networks and data centers; cloud platform architecture with security over virtualized data centers; service-oriented architectures for distributed computing; and cloud programming and software environments. Additionally, students will learn how to conduct some parallel and distributed programming and performance evaluation experiments on applications within available cloud platforms. Finally we will survey research literature and latest technology trends that are shaping the future of high performance, distributed and cloud computing. Graduate students enrolled in the course will have a collaborative programming project using tools and software environments available within real cloud platforms.

CMP_SC 7001 (Topic) – Big Data Analytics (Plan to move to online format)
This course will cover advanced analytics technologies and techniques that enable industries to extract information from data with previously unachievable levels of sophistication, speed and accuracy. Students will learn practical industry best practices using a current big data analytical platform to bridge the gap between classroom learning and real world applications.

INFOINST 8750 – GPU and Multi-Core Computing (to be developed)
This course examines the use of special-purpose hardware originally designed for graphics and games to solve general-purpose computing problems. Graphics processing units (GPUs) have high peak performance for arithmetically-intensive computations, and a lower cost than their general-purpose counterparts with similar performance levels. Students in the course will learn how to develop scalable parallel programs targeting the unique requirements for obtaining high performance on GPUs. The course will contrast parallel programming for GPUs with parallel programming for conventional multi-core microprocessors.

Strategic Communications and Data Journalism Emphasis Area Courses

JOURN 7263 – Interacting Advertising II (Already offered online)
Goes in-depth on top issues in the interactive process from video advertising to social networking sites and how to increase campaign performance with web analytics. Designed for those who want a career in interactive advertising.

JOURN 7236 – Psychology of Advertising
The strategic communication industry is abandoning an outdated and flawed view of consumers based solely on consumer behavior in favor of a cutting edge approach grounded in neuroscience called Neuromarketing. This approach has been recognized by the marketing research industry as one of the fastest-growing segments of marketing communications research and every major global advertiser is currently either conducting or systematically investigating the use of this kind of research. The promise of this new form of strategic communication science is that it provides a more scientifically valid view of how consumers actually process communications and make decisions by measuring less conscious brain processes.

JOURN 7248 – Media Strategy Planning
This course will introduce the wild and wonderful world of media planning. It is an area that is rapidly evolving with new forms of media and measurement. Students will learn terms used in the industry, standard formulas for calculating media value, the process for developing a media strategy and how to select tactics. Students will be introduced to working professionals and exposed to industry resources commonly used by practitioners. They will explore media topics through individual and group work. This course will also give students the opportunity to practice and polish their writing and presentation skills which are critical to nearly every aspect of advertising.

JOUR 7430 – Computer-Assisted Reporting (need to change some content and adapt for online)
Learn how to identify, obtain, evaluate, clean and analyze data, primarily from publicly-available sources. Students will learn how to turn their data and discoveries into news stories, information graphics or interactive news applications. In addition, students will be expected to successfully negotiate for data from government sources.

JOUR 7430 – Advanced Data Journalism (need to change some content and adapt for online)
The goal of the class is to teach students how to solve problems in journalism using computer programming. Students will use the python language and Django web framework. At the end of the semester, students will have the skills to practice data journalism for a news organization.

JOUR 7440 – Mapping for Stories and Graphics (need adapt for online and expand to 3 cr.)
Learn how to use geographic data in data journalism. The primary focus is on using geographic information systems (GIS) to analyze data spatially for news stories. Also, students will learn how to produce data that can be used for online visualizations.

Human Centered Data Science Design
IS_LT 9410 (Topic) – Introduction to Human Centered Data Science Design (already offered, mixed-mode)
Human Centered Data Science Design combines both the technical (mathematical modeling, databases, social networking, and text mining) and social (economic, ethical, policy, and political) aspects of data analytics. Students work through a series of case studies with real data to develop both the theoretical and hands-on experience necessary to fill leadership roles in e-science, eResearch, and big data.

IS_LT 7310 – Seminar in Human-Computer Interaction (already offered; to be modified for mixed-mode)
Increasingly human activity is supported, mediated and enacted through various forms of computing. Web searches replace visits to the library, ecommerce replaces trips to the mall, and e-learning replaces going to class. Email, chat and instant messaging supplement social interaction. Calculators, spell checkers, and other tools support improved human performance. Networked games and digital music are key sources of entertainment and relaxation. Workplaces are augmented, automated and amplified by information systems. Through readings, demonstrations and two research projects students will explore how humans interact with computers and how interactive technology augments and transforms human activity. Students will build knowledge of human-computer interaction (HCI), the field and specialties that advance HCI, and the methods used to design interfaces and study effects of HCI.

ISLT 9410 (Topic) – Learning Analytics (already offered mixed-mode delivery)
This course on learning analytics will focus on the development of adaptive and adaptable indicators of learning, subject matter interest, and social engagement in various technologically mediated learning environments. The emphasis will be placed on the collection, identification, management, analysis and visualization of behavioral data from formal and informal learning environments. Examples of the learning environments and student outcomes measured include Virtual Math Teams, iSocial, Sakai, GitHub, Moodle and other technologically mediated environments where learning occurs. Students will be introduced to social network analysis (SNA), Group Informatics, computational linguistics and Bayesian statistical methods.

Digital Humanities option in Human Centered Data Science Design
Students May select One course from this sequence in place of the Learning Analytics or Seminar in Human Computer Interaction

IS_LT 9410 (Topic) – Digital Humanities and Information (Already offered, mixed-mode delivery)
This course on history, philosophy, and methods in the emerging field of 'digital humanities' (DH) focuses on topics at the intersection of information, the humanities disciplines, technology, and culture as well as the contexts of the academy, libraries, archives, museums, and media. Case studies from a range of disciplines highlight the debates, and show how digital humanities techniques combined with interdisciplinary approaches from the fields of information and media offer new ways to look at old problems, and open new research avenues that are not possible using traditional means.

IS_LT 9410 (Topic) – Metadata (Already offered online)
Explores principles, standards, and schema for metadata in diverse online environments to facilitate retrieval and sharing within and between information systems. Emphasizes the creation and use of metadata for specific purposes by various communities.

IS_LT 9410 (Topic) – *Learning Analytics (already offered, mixed-mode delivery)*
This course on learning analytics will focus on the development of adaptive and adaptable indicators of learning, subject matter interest, and social engagement in various technologically mediated learning environments. The emphasis will be placed on the collection, identification, management, analysis and visualization of behavioral data from formal and informal learning environments. Examples of the learning environments and student outcomes measured include Virtual Math Teams, iSocial, Sakai, GitHub, Moodle and other technologically mediated environments where learning occurs. Students will be introduced to social network analysis (SNA), Group Informatics, computational linguistics and Bayesian statistical methods.

Appendix B: Support Letters

The following letters of support are attached to this proposal:

- Chase Davis, Deputy Editor, Interactive News – *The New York Times*
- Don Etling, Senior VP/Senior Partner – *Fleishman Hillard*
- Gary Allen, VP for Information Technology and MU CIO – *University of Missouri System*
- Glen Schuster. Sr. VP and CTO – *Centene*
- Harlen D. Hays, Sr. Manager, Qualitative Research and Biostatistics – *Cerner*
- Heather Nelson, Solution Architect – *Silicon Valley Data Science*
- John Everson, Director Information Security – *Dish Network*
- John Lee, Director, Data Analytics – *KPMG*
- Paul Pancoast – *Deloitte Consulting*
- Steve Rounsley, Genomics Lead, Seeds and Traits R&D – *Dow AgroSciences*
- Tom Henry, Chief Data Officer & VP – *Express Scripts*
- Valinda Kennedy, IBM Chicago Cloud Academic, Entrepreneur and Developer Gladiator – *IBM*

Appendix C: Budget and Financial Projections