THE UNIVERSITY OF ARIZONA

New Academic Program Workflow Form

General

Proposed Name: Data Science

Transaction Nbr: 00000000000089

Plan Type: Major

Academic Career: Graduate

Degree Offered: Master of Science

Do you want to offer a minor? N

Anticipated 1st Admission Term: Fall 2021

Details

Department(s):

SBSC

DEPTMNT ID	DEPARTMENT NAME	HOST
0481	School of Information	Y

Campus(es):

ONLN

LOCATION	DESCRIPTION
ONLN	Online

Admission application terms for this plan: Spring: Y Summer: Y Fall: Y

Plan admission types:

Freshman: N Transfer: N Readmit: N Graduate: Y

Non Degree Certificate (UCRT only): Y

Other (For Community Campus specifics): N

Plan Taxonomy: 30.7001, Data Science, General.

Program Length Type: Program Length Value: 0.00

Report as NSC Program:

SULA Special Program:

Print Option:

Diploma: Y Master of Science in Data Science

Transcript: Y Master of Science in Data Science

Conditions for Admission/Declaration for this Major:

The application will have 4 explicit questions asking the applicant to list coursework or professional experience that meets these criteria.
1. Strong quantitative and analytical reasoning abilities, as demonstrated by academic coursework and/or professional experience.
2. Experience with math and programming, including data structures, analysis of algorithms, and linear algebra, as demonstrated by academic coursework and/or professional experience.
3. What programming language do you use most frequently?
4. In the past 3 months, how much time (on average) have you spent each week coding?

Requirements for Accreditation:

N/A

Program Comparisons

University Appropriateness

Data Science is part of the landscape of the fourth industrial revolution. UA's Strategic Plan focuses on transdisciplinary convergence, to realize the transformative power of emerging data science methodologies and tools. The School of Information's proposed MS degree presents a compelling opportunity to build on UA's unique strengths in interdisciplinary efforts given the existence of data science across units and programs. This degree advances the College of Social and Behavioral Sciences strategic plan to offer a broad-based education and the iSchool's focus on issues that lay at the intersections of data, technology, and people.

Arizona University System

NBR	PROGRAM	DEGREE	#STDNTS	LOCATION	ACCRDT
1	Information and Data Science	MS	792	UC Berkeley	N
2	Data Science	MS	393	University of Wisconsin	N

Peer Comparison

See attached chart

Faculty & Resources

Faculty

Current Faculty:

INSTR ID	NAME	DEPT	RANK	DEGREE	FCLTY/%
00202612	Bruce Fulton	0481	Assoc. Prof	Doctor of Philosophy	1.00
08308838	John Hartman	0412	Assoc. Prof	Doctor of Philosophy	1.00
11709208	Steven Bethard	0481	Assoc. Prof	Doctor of Philosophy	1.00
12200140	Young-Jun Son	2302	Professor	Doctor of Philosophy	1.00
16508028	Hong Cui	0481	Professor	Doctor of Philosophy	1.00
17206152	Clayton Morrison	0481	Assoc. Prof	Doctor of Philosophy	.60
17905279	Patrick Heidorn	0481	Professor	Doctor of Philosophy	1.00
22052443	Peter Jansen	0481	Assit. Prof	Doctor of Philosophy	1.00
22067613	Gregory Ditzler	2303	Assit. Prof	Doctor of Philosophy	1.00
22083799	Jason Pacheco	0412	Assit. Prof	Doctor of Philosophy	1.00
22083805	Chicheng Zhang	0412	Assit. Prof	Doctor of Philosophy	1.00
22084317	Ryan Rucker	0481	Adj. Instor.	Doctor of Education	1.00
22088136	Meaghan Wetherell	0481	Assit. Prof	Doctor of Philosophy	1.00
23152438	Gustave Hahn-Powell	0431	Assit. Prof	Doctor of Philosophy	1.00
23263591	Adriana Picoral Sarandy Machado Scheid	0481	Assit. Prof	Doctor of Philosophy	1.00

Additional Faculty:

None

Current Student & Faculty FTE

DEPARTMENT	UGRD HEAD COUNT	GRAD HEAD COUNT	FACULTY FTE
0481	634	250	44.78

Projected Student & Faculty FTE

	UGRD HEAD COUNT			GRAD HEAD COUNT			FACULTY FTE		
DEPT	YR 1	YR 2	YR 3	YR 1	YR 2	YR 3	YR 1	YR 2	YR 3
0481	648	669	690	293	349	449	44.78	44.78	44.78

Library

Acquisitions Needed:

None

Physical Facilities & Equipment

Existing Physical Facilities:

Program will be online and will not need physical facilities

Additional Facilities Required & Anticipated:

Program will be online and will not need physical facilities

Other Support

Other Support Currently Available:

None

Other Support Needed over the Next Three Years:

None

Comments During Approval Process

2/1/2021 1:58 PM

CFBROOKS

Comments
Approved.

2/1/2021 2:02 PM

RICAR22

Comments
Approved.

2/1/2021 2:36 PM

CRAIGWILSON

Comments
Approved.

3/5/2021 11:02 AM

ESANDMAR

Comments
Uploaded updated add'l info form.

THE UNIVERSITY OF ARIZONA

NEW ACADEMIC PROGRAM-GRADUATE MAJOR

ADDITIONAL INFORMATION FORM

I. MAJOR DESCRIPTION -provide a marketing/promotional description for the proposed program. Include the purpose, nature, and highlights of the curriculum, faculty expertise, emphases (sub-plans; if any), etc. The description should match departmental and college websites, Graduate Catalog and Program Descriptions page, handouts, promotional materials, etc.

The theory and tools behind Data Science are used throughout the academy, stretching from the humanities to the biological sciences and from the fine arts to physics and engineering. Data Science is thus both transdisciplinary across the academy and at the same time foundational in many specific fields. The 30-unit M. S. program in Data Science in the iSchool draws upon the expertise of iSchool faculty directly engaged in issues at the intersection of technology, people, and information, thus drawing upon a central foundation in information science.

However, by making use of the novel structure of stacked certificates, it also engages with programs throughout the university to provide students with opportunities to develop skills and knowledge appropriate to their specific interests and domains of study.

The stacked certificate model constructs an online master's degree, where students enroll in a core information science certificate, and then also pursue elective certificates in a wide range of specializations. They combine these certificates into a single Master's degree, which is completed by a synthesizing capstone project or portfolio that integrates the three certificates.

The stacked certificate model allows students a variety of options for entry into the degree. If they are uncertain that they want to purse a full MS they can take any of the certificates as free-standing options. If they decide they want the MS they can apply at any point (within 4 years of the completion of a certificate) and synthesize their experiences through the capstone. Other students may wish to simply enroll directly into the MS degree from the beginning and take the certificates en route to their MS.

The MS degree is structured as follows:

1) Required Core Certificate (9 units) offered by the iSchool
2) A required synthesizing capstone project (3 units) managed by the iSchool along with campus partners.
3) Option to select certificates from this list (9 units each), future certificates (9-15 units, typically) developed from across domain areas at the University of Arizona, or to choose a set of elective course choices to complete their plan.

Available Elective Certificates for Academic Year 2021/2022
Certificate 1: Foundations of Data Science (School of Information)
Certificate 2: Natural Language Processing (Linguistics)

Planned certificates for Spring 2022 launch
Certificate 3: Database Management (School of Information)
Certificate 4: Spatial Data Science (School of Geography)

We are in negotiations to develop additional elective certificates in a wide variety of disciplines with other programs and colleges. Additional elective certificate proposals are always welcome as are the option to include existing certificates in the area of Data Science.

Through transdisciplinary stacked certificate curriculum, the MS in Data Science program produces generations of professionals and researchers who have the knowledge, skill, and ability to engage in data science processes and work. This degree will encompass a blend of disciplines to address the grand challenges of today and the future. The MS in Data Science provides students the training they need in data collection, exploration, manipulation and storage, analysis, and presentation in order to navigate data-rich workplace environments.

CIP CODE – 30.7001 Data Science, General.

II. NEED FOR THE MAJOR/JUSTIFICATION-describe how the major fulfills the needs of the city, state, region, and nation. Provide market analysis data or other tangible evidence of the need for and interest in the proposed major (and emphases, if applicable). This might include results from surveys of current students, alumni, and/or employers or reference to student enrollments in similar programs in the state or region. Include an assessment of the employment opportunities for graduates of the program for the next three years. Curricular Affairs can provide a job posting/demand report by skills obtained/outcomes/CIP code of the proposed major. Please contact the Office of Curricular Affairs to request the report for your proposal.

The proposed MS in Data Science program housed in the iSchool will support the priorities outlined in the UA's Strategic Plan. This degree will prepare graduates to be innovative scientific leaders. Enhancing the supply of well-trained leaders in the data sciences will help catalyze economic development in the state of Arizona. Built upon the foundations of probability theory, optimization, as well as algorithmic thinking and doing, the data sciences include a broad set of approaches to include bioinformatics, natural language processing, machine learning, data mining, data visualization, and more. The stacked certificate model means that the program can provide a structure for transdisciplinary collaborations.

The state of Arizona's economy relies heavily on technological design and information-based industries. According to the Arizona Commerce Authority (*Arizona Business Know How*), the core of Arizona's competitiveness plan includes Aerospace & Defense as well as Technology & Innovation. Arizona seeks to foster deep partnerships with major electronics (Intel, Motorola, IBM, Raytheon, etc.) and biotechnology (Ventana, Medtronic, Flinn Foundation, etc.) industries. To support these industries, the State needs a workforce that is broadly educated in data science. This degree program is unique in its combination of training in the foundations of data science, while also training learners to be inherently interdisciplinary.

According to U.S. News, Data Scientist jobs are 8 out of 100 best jobs in the United States. With a median salary of $94,280 and a 3.5% unemployment rate, this newer career field is increasingly recognized by employers as a valuable asset to their companies (*Data Scientist Ranks Among Best Jobs of 2021*, n.d.) The Bureau of Labor Statistics projects 31 percent employment growth for data scientists between 2019 and 2029 (*Fastest Growing Occupations: Occupational Outlook Handbook,* 2020). The online Master's in Data Science at the UC Berkeley School of Information has graduates working in jobs such as data analyst, data architect, data engineer, solutions architect, and systems engineer at top tier companies like Amazon, Apple, Facebook, Microsoft, and Google (*Online Master of Data Science Degree*, n.d.). The Master's in Data Science program at the Indiana University

School of Information is also exemplary, and covers most primarily the domains of data analytics and visualization, intelligent systems engineering, precision health, and cybersecurity (*M.S. in Data Science—Online*, n.d.).

For iSchools, data science programs involve faculty and courses across disciplines and organizational units. iSchools are inherently interdisciplinary places given the broad reach of informatics and information science. iSchools "…have been newly created or are evolving from programs formerly focused on specific tracks such as information technology, library science, informatics, and information science. While each individual iSchool has its own strengths and specializations, together they share a fundamental interest in the relationships between information, people, and technology" (*https://ischools.org/About*). Schools are distinct in their working across disciplines, always focused on making data findable, visible and useful for people across domains. At the University of Arizona, as well as on other campuses, data science resides in a wide variety of units and Colleges. Almost every College on Campus has a program that fits within the broad definition of Data Science. So, units and Colleges across the campus are invited to participate in this degree by adding certificates that can be 'stacked' into this M. S. program.

References

Data Scientist Ranks Among Best Jobs of 2021. (n.d.). Retrieved January 12, 2021, from https://money.usnews.com/careers/best-jobs/data-scientist.

Fastest Growing Occupations: Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics. (September 1, 2020). Retrieved January 12, 2021, from https://www.bls.gov/ooh/fastest-growing.htm.

M.S. in Data Science—Online. (n.d.). Data Science Program. Retrieved January 12, 2021, from https://datascience.indiana.edu/programs/ms-data-science-online/index.html.

Online Master of Data Science Degree. (n.d.). Retrieved January 12, 2021, from https://ischoolonline.berkeley.edu/data-science/.

The iSchools Organization. (n.d.). Retrieved April 7, 2021 from https://ischools.org/About.

III. MAJOR REQUIREMENTS– complete the table below by listing the major requirements, including required number of units, required core, electives, and any special requirements, including emphases (sub-plans), thesis, internships, etc.

Total units required to complete the degree	30
Pre-admissions expectations (i.e. academic training to be completed prior to admission)	Applicants are expected to have completed undergraduate coursework in programming and statistics. Coursework in calculus preferred but not required. The application will have 2 explicit questions asking the applicant to list coursework or professional experience that meets these criteria. 1. What academic coursework and/or professional experience demonstrates your strong quantitative and analytical reasoning abilities? 2. What academic coursework and/or professional experience demonstrates your experience with math and programming, including data structures, analysis of algorithms, and linear algebra?
Major requirements. List all major requirements including core and electives. If applicable, list the emphasis requirements for each proposed emphasis*. Courses listed must include course prefix, number, units, and title. Mark new coursework (New). Include any limits/restrictions needed (house number limit, etc.). Provide email(s)/letter(s) of support from home department head(s) for courses not owned by your department.	Core Courses, 9 units (comprise cert. #1 Foundations of Data Science) INFO 520 (3 units) Ethical Issues in Information INFO 523 (3 units) Data Mining and Discovery INFO 526 (3 units) Data Visualization Capstone – one course, 3 units INFO 698 Capstone Project Elective Courses, 18 units (option to choose two certificates of 9 units each or 18 units from across the entire set of elective courses). *Elective Courses Organized as Optional Certificates* Certificate 2: Natural Language Processing (Linguistics) LING 539 (3 units) Statistical Natural Language Processing LING 582 (3 units) Advanced Statistical Natural Language Processing *Plus one more course chosen from:*

	LING 578 (3 units) Speech Technology LING 581 (3 units) Advanced Computational Linguistics INFO 557 (3 units) Neural Networks CSC 583 (3 units) Text Retrieval and Web Search CSC 585 (3 units) Algorithms for Natural Language Processing Certificate 3: Database Management (School of Information and School of Geography) INFO 579: (3 units) Database Design in SQL INFO 570: (3 units) Database Development and Management GIST 570: (3 units) Geodatabase Certificate 4: Geospatial Data Science (School of Geography) GIST XXX: (3 units) Geospatial Visualization GIST 603A: (3 units) Geog. Info. Systems Programming/Automation GIST 604B: (3 units) Open Source GIS *Elective Courses Not Organized as part of Certificates* (if not opting for a set of certificates, courses can also be chosen from this set). ECE 523 (3 units) Machine Learning ECE 524 (3 units) Fundamentals of Cloud Security ECE 579 (3 units) Artificial Intelligence INFO 514/POL 514 (3 units) Computational Social Science INFO 521 (3 units) Machine Learning INFO 531 (3 units) Data Warehousing and Analytics in the Cloud INFO 536 (3 units) Data Science and Public Interests INFO 555 (4 units) Applied Natural Language Processing INFO 556 (3 units) Text Retrieval and Web Search INFO 578 (3 units) Science Information and its Presentation SIE 530: (3 units) Engineering Statistics SIE 533: (3 units) Fundamentals of Data Science for Engineers SIE 545 (3 units) Fundamentals of Optimization SIE 640 (3 units) Large-Scale Optimization SIE 645 (3 units) Nonlinear Optimization
Research methods, data analysis, and methodology requirements (Yes/No). If yes, provide description.	Yes, research methods, data analysis, and methodology are taught in three of the four core courses as well as in all of the elective courses.

Internship, practicum, applied course requirements (Yes/No). If yes, provide description.	Yes, Capstone course. INFO 698 Capstone Project
Master thesis or dissertation required (Yes/No). If yes, provide description.	No
Additional requirements (provide description)	None
Minor options (as relevant)	

IV. CURRENT COURSES– using the table below, list all existing courses included in the proposed major. You can find information to complete the table using the UA course catalog or UAnalytics (Catalog and Schedule Dashboard> "Printable Course Descriptions by Department" On Demand Report; right side of screen). If the courses listed belong to a department that is not a signed party to this implementation request, upload the department head's permission to include the courses in the proposed program and information regarding accessibility to and frequency of offerings for the course(s). Upload letters of support/emails from department heads to the "Letter(s) of Support" field on the UAccess workflow form. Add rows to the table, as needed.

Course prefix and number	Units	Title	Course Description	Pre-reqs	Modes of delivery	Typically Offered (F,W, Sp, Su)
CSC 583	3	Text Retrieval and Web Search	Most of the web data today consists of unstructured text. Of course, the fact that this data exists is irrelevant, unless it is made available such that users can quickly find information that is relevant for their needs. This course will cover the fundamental knowledge necessary to build these	MATH 129	Online, in-person	F

			systems, such as web crawling, index construction and compression, Boolean, vector-based, and probabilistic retrieval models, text classification and clustering, link analysis algorithms such as PageRank, and computational advertising. The students will also complete one programming project, in which they will construct one complex application that combines multiple algorithms into a system that solves real-world problems.			
CSC 585	3	Algorithms for Natural Language Processing	This course covers important algorithms useful for natural language processing (NLP), including distributional similarity algorithms such as word embeddings, recurrent and recursive neural networks (NN), probabilistic graphical models useful for sequence prediction, and parsing algorithms such as shift-reduce. This course will focus on the algorithms that underlie NLP, rather than the application of NLP to various problem domains.	None	Online, in-person	F, Sp
ECE 523	3	Engineering Applications of Machine Learning and Data Analytics	Machine learning deals with the automated classification, identification, and/or characterizations of an unknown system and its parameters. There are an overwhelming number of application driven fields that can benefit from machine learning techniques. This course will introduce you to machine learning and develop core principles that allow you to determine which algorithm to use, or design a novel approach to solving to engineering	ECE 503	Online, in-person	Sp

			task at hand. This course will also use software technology to supplement the theory learned in the class with applications using real-world data.			
ECE 524	3	Fundamentals of Cloud Security	Cloud Computing is an emerging paradigm that aims at delivering computing, information services, and data storage as a utility service over a network (e.g., Internet). There is a strong interest in cloud computing due to their performance and host, but their rapid deployment will exacerbate the security problem. In cloud computing, organizations relinquish direct control of many security aspects to the service providers such as trust, privacy preservation, identity management, data and software isolation, and service availability. The adoption and proliferation of cloud computing and services will be severely impacted if cloud security is not adequately addressed. The main goal of this course is discuss the limitations of current cybersecurity approaches to clouds and then focus on the fundamental issues to address the cloud security and privacy such as the confidentiality, integrity and availability of data and computations in clouds. In this course we will examine cloud computing models, look into the threat model and security issues related to data and computations outsourcing, and explore practical applications to make cloud resources secure and resilient to cyber attacks.	None	Online, in-person	Sp

ECE 579	3	Principles of Artificial Intelligence	Provides an introduction to problems and techniques of artificial intelligence (AI). Automated problem solving, methods and techniques; search and game strategies, knowledge representation using predicate logic; structured representations of knowledge; automatic theorem proving, system entity structures, frames and scripts; robotic planning; expert systems; implementing AI systems.	None	Online, in-person	Sp
GIST 570	3	Geo-Databases	This course introduces essentials of database design, development, and analysis for GIS. Emphasis is on geospatial data and appropriate database designs and implementations for GIS Enterprise. Topics include: requirements engineering for GIS data, business case design using the Entity-Relationship model, object-relational database design, database normalization, database optimization, data handling, security and risk management, and IT auditing. Database technologies will be demonstrated with Spatial Database Management Systems such as PostgreSQL/PostGIS, Oracle, and ArcGIS Server/Enterprise. Server programming will make use of SQL and Procedural SQL in PostGIS. A business case will be developed to train the student in the database lifecycle supporting organizational operations, planning, and data management in GIS.	None	Online, in-person	F, Sp

GIST 603A	3	Geographic Information Systems Programming and Automation	The goal of this course is to gain an introductory understanding of geographic programming and data automation techniques using ModelBuilder and the Python language.	None	Online, in-person	F, Sp, Su
GIST 604B	3	Open Source GIS	The focus of this class is to examine and apply GIS open source programming. We will examine common languages used like Python, Java, html 5, as well as APIs, JSON, html, and SQL, to automate workflows, extend the tools, and create interactive web and mobile GS platforms. Topics include preparing data as strings, lists, tuples, and dictionaries prior to use, using Python to run SQL queries, working with roasters in Python, automating mapping tasks, and developing custom scripting tools. In addition to weekly assignments and readings, assessment will be oriented around a single, student-directed project that will take the second half of the semester to complete. It will require students to write a simple script to accomplish a specified task in ArcGIS and present the results of their work to peers.	None	Online, in-person	F, SP
INFO 514 (POL514)	3	Computational Social Science	This course will guide students through advanced applications of computational methods for social science research. Students will be encouraged to consider social problems from across sectors, like health science, education, environmental policy and business. Particular attention will be given to	None	Online, in-person	Sp

			the collection and use of data to study social networks, online communities, electronic commerce and digital marketing. Students will consider the many research designs used in contemporary social research and will learn to think critically about claims of causality, mechanisms, and generalization in big data studies. Graduate requirements include additional readings and a more in-depth final paper than is required at the undergraduate level.			
INFO 520	3	Ethical Issues in Information	This course presents an overview and understanding of the intractable and pressing ethical issues as well as related policies in the information fields. Emerging technological developments in relation to public interests and individual well-being are highlighted throughout the course. Special emphasis is placed on case studies and outcomes as well as frameworks for ethical decision-making.	INFO 420	Online, in-person	F, Sp
INFO 521	3	Introduction to Machine Learning	Machine learning describes the development of algorithms which can modify their internal parameters (i.e., "learn") to recognize patterns and make decisions based on example data. These examples can be provided by a human, or they can be gathered automatically as part of the learning algorithm itself. This course will introduce the fundamentals of machine learning, will describe how to implement several practical methods for pattern	Must have taken ISTA 311, MATH 129, AND MATH 313, or equivalent, or consen	Online, in-person	F

			recognition, feature selection, clustering, and decision making for reward maximization, and will provide a foundation for the development of new machine learning algorithms.	t of instructor. ISTA 116 or comparable is recommended.		
INFO 523	3	Data Mining and Discovery	This course will introduce students to the concepts and techniques of data mining for knowledge discovery. It includes methods developed in the fields of statistics, large-scale data analytics, machine learning, pattern recognition, database technology and artificial intelligence for automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns. Topics include understanding varieties of data, data preprocessing, classification, association and correlation rule analysis, cluster analysis, outlier detection, and data mining trends and research frontiers. We will use software packages for data mining, explaining the underlying algorithms and their use and limitations. The course include laboratory exercises, with data mining case studies using data from many different resources such as social networks, linguistics, geo-spatial applications, marketing and/or psychology	None	Online, in-person	F, Sp

INFO 531	3	Data Warehousing and Analytics in the Cloud	Data Warehousing and Analytics In the Cloud will utilize concepts, frameworks, and best practices for designing a cloud-based data warehousing solution and explore how to use analytical tools to perform analysis on your data. In the first half of the course, I will provide an overview of the field of Cloud Computing, its main concepts, and students will get hands-on experience through projects which utilize cloud computing platforms. In the second half of the course, we will examine the construction of a cloud-based data warehouse system and explore how the Cloud opens up data analytics to huge volumes of data.	Familiarity with database concepts and basic SQL syntax	Online, in-person	F, Sp, Su
INFO 536	3	Data Science and Public Interest	This course focuses on the use of modern data science methods to help learners make socially responsible decisions and mitigate harm that arises from issues like bias, discrimination, and threats to one's personal privacy. More and more individuals are needing to make data-driven decisions in a wide variety of contexts including non-governmental organizations, not-for-profit industries, human services, environmental organizations, refugee camps, and more. Students in this class will thus learn about data science and how it can be utilized in contexts where socially-good decisions are desired and emphasized. This active learning class is designed for students who	None	Online, in-person	F, Sp

			have an interest in the topic but who may have little to no previous experience with data science or programming.			
INFO 539 (LING 539, CSC 539)	3	Statistical Natural Language Processing	This course introduces the key concepts underlying statistical natural language processing. Students will learn a variety of techniques for the computational modeling of natural language, including: n-gram models, smoothing, Hidden Markov models, Bayesian Inference, Expectation Maximization, Viterbi, Inside-Outside Algorithm for Probabilistic Context-Free Grammars, and higher-order language models. Graduate-level requirements include assignments of greater scope than undergraduate assignments. In addition to being more in-depth, graduate assignments are typically longer and additional readings are required.	LING 538	Online, in-person	F
INFO 555	4	Applied Natural Language Processing	Most of web data today consists of unstructured text. This course will cover the fundamental knowledge necessary to organize such texts, search them a meaningful way, and extract relevant information from them. This course will teach natural language processing through the design and development of end-to-end natural language understanding applications, including sentiment analysis (e.g., is this review positive or negative?), information extraction (e.g., extracting named entities and their relations from text), and question answering (retrieving exact answers to natural language	ISTA 455	Online, in-person	Sp

			questions such as "What is the capital of France" from large document collections). We will use several natural language processing toolkits, such as NLTK and Stanford's CoreNLP. The main programming language used in the course will be Python, but code written in Java or Scala will be accepted as well. Graduate-level requirements include implementing more complex, state-of-the-art algorithms for the three proposed projects. This will require additional reading of conference papers and journal articles.			
INFO 556	3	Text Retrieval and Web Search	Most of the web data today consists of unstructured text. Of course, the fact that this data exists is irrelevant, unless it is made available such that users can quickly find information that is relevant for their needs. This course will cover the fundamental knowledge necessary to build such systems, such as web crawling, index construction and compression, boolean, vector-based, and probabilistic retrieval models, text classification and clustering, link analysis algorithms such as PageRank, and computational advertising. The students will also complete one programming project, in which they will construct one complex application that combines multiple algorithms into a system that solves real-world problems. Graduate level requirements include implementing more complex, state-of-the-art algorithms for the programming project, which might require additional reading of research articles. Written	ISTA 456	Online, in-person	F

			assignments will have additional questions for graduate students.			
INFO 557	3	Neural Networks	Neural networks are a branch of machine learning that combines many simple computational units to allow computers to learn from and generalize over complex patterns in data. Students in this course will learn how to train and optimize feed forward, convolutional, and recurrent neural networks for tasks such as text classification, image recognition, and game playing.	Basic programming skills and some experience with analysis of algorithms and data structures. Basic linear algebra skills recommended.	Online, in-person	Sp
INFO 570	3	Database Development and Management	This course covers theory, methods, and techniques widely used to design and develop a relational database system and students will develop a broad understanding of modern database management systems. applications of fundamental database principles in a stand-alone database environment using MS Access and Windows are emphasized. Applications in an Internet environment will be discussed using MySQL in the Linux platform. Graduate-level requirements include a group project	None	Online, in-person	F, Sp

			consisting of seven sections: Database Design; Implementation (Tables); Forms; Data Retrieval (Queries/Reports); Project Presentation; Project Report; and, Peer Evaluation.			
INFO 578 (LIS 578)	3	Science Information and its Presentation	In today's digital society, people have access to a wide variety of information sources and scientific data. In this course, students will learn about the role of science and scientific data in society, and they will consider means for making science information findable and understandable for a wide variety of audiences. This course will provide students an interdisciplinary experience for considering science data and how that information gets shared across contexts.	None	Online, in-person	Sp
INFO 698	3	Capstone Project	Capstone Project is intended to provide an opportunity for students to show off what they have mastered in the program. The project should be relevant to MS degree competencies and contribute to the development and enforcement of the student's knowledge and skill sets in the field of Information Science. The student should propose a project plan and the faculty advisor should approve it before registration. The project plan should include goals for the project, MS competencies addressed by the project, system design, an implementation schedule, and the assessment plan. The project plan should also include reasonable milestones and check points. The amount of the work		Online, in-person	F, Sp, Su

			should be appropriate for a 3-unit course. The primary faculty advisor must be an SI faculty, but faculty members from other units may participate in advising the student.			
LING 578	3	Speech Technology	Topics include speech synthesis, speech recognition, and other speech technologies. This course gives students background for a career in the speech technology industry. Graduate students will do extra readings, extra assignments, and have an extra presentation. Their final project must constitute original work in a speech technology.		Online, in-person	Sp
LING 581	3	Advanced Computational Linguistics	This course provides a hands-on project-based approach to particular problems and issues in computational linguistics.	LING 538	Online, in-person	Sp
LING 582	3	Advanced Statistical Natural Language Processing	This course focuses on statistical approaches to pattern classification and applications of natural language processing to real-world problems	LING 539	Online, in-person	F
SIE 530	3	Engineering Statistics	Statistical methodology of estimation, testing hypotheses, goodness-of-fit, nonparametric methods and decision theory as it relates to engineering practice. Significant emphasis on the underlying statistical modeling and assumptions. Graduate-level requirements include additionally		Online, in-person	F

			more difficult homework assignments.			
SIE 533	3	Fundamentals of Data Science for Engineers	This course will provide senior undergraduate and graduate students from a diverse engineering disciplines with fundamental concepts, principles and tools to extract and generalize knowledge from data. Students will acquire an integrated set of skills spanning data processing, statistics and machine learning, along with a good understanding of the synthesis of these skills and their applications to solving problem. The course is composed of a systematic introduction of the fundamental topics of data science study, including: (1) principles of data processing and representation, (2) theoretical basis and advances in data science, (3) modeling and algorithms, and (4) evaluation mechanisms. The emphasis in the treatment of these topics will be given to the breadth, rather than the depth. Real-world engineering problems and data will be used as examples to illustrate and demonstrate the advantages and disadvantages of different algorithms and compare their effectiveness as well as efficiency, and help students to understand and identify the circumstances under which the algorithms are most appropriate.	SIE 530, equivalent courses such as SIE 500A taken in parallel or consent of instructor.	Online, in-person	F
SIE 545	3	Fundamentals of Optimization	Unconstrained and constrained optimization problems from a numerical standpoint. Topics include variable metric methods, optimality	SIE 340	Online, in-person	F

			conditions, quadratic programming, penalty and barrier function methods, interior point methods, successive quadratic programming methods.			
SIE 640	3	Integer and Combinatorial	Modeling and solving problems where the decisions form a discrete set. Topics include model development, branch and bound methods, cutting plane methods, relaxations, computational complexity, and solving well-structured problems.	SIE 544	Online, in-person	Sp
SIE 645	3	Nonlinear Optimization	This course is devoted to structure and properties of practical algorithms for unconstrained and constrained nonlinear optimization.	SIE 544 or SIE 5454	Online, in-person	Sp

V. NEW COURSES NEEDED – using the table below, list any new courses that must be created for the proposed program. If the specific course number is undetermined, please provide level (ie CHEM 6**). Add rows as needed. Is a new prefix needed? If so, provide the subject description so Curricular Affairs can generate proposed prefix options.

Course prefix and number	Units	Title	Course Description	Pre-req	Modes of delivery	Typically Offered
INFO 526 Under Curr. Review (since Jan 2021)	3	Data Visualization	This course provides an overview of the various concepts and skills required for effective data visualization. It presents principles of graphic design, programming skills, and statistical knowledge required to build compelling visualizations that communicate effectively to target audiences. Visualization skills addressed in this course include choosing appropriate colors, shapes, variable mappings, and interactivity based on principles of color perception, pre-attentive processing, and accessibility.	None	Online, in-person	F, Sp
INFO 579 In Dev. (D)	3	Database Design in SQL	The course provides students with training in writing basic queries in the structured query language (SQL). Upon completion of this course, students will have an	None	Online, in-person	F, Sp

			advanced understanding of database design in SQL.			
GIST XXX	3	Geospatial Visualization	Description in progress.	None	Online, in-person	F, Sp

VI. FACULTY INFORMATION- complete the table below. If UA Vitae link is not provided/available, attach a short CV (2-3 pages) to the end of the proposal or upload to the workflow form (in the "Letter(s) of Support" field). UA Vitae profiles can be found in the UA directory/phonebook.

Faculty Member	Involvement	UA Vitae link or "CV attached"
Steven Bethard	iSchool, Teaches NLP, Neural Networks, and other data science courses	https://profiles.arizona.edu/person/bethard
Hong Cui	iSchool, Teaches INFO 523, Data Mining	https://profiles.arizona.edu/person/hongcui
Gregory Ditzler	ECE, Teaches data science courses	https://profiles.arizona.edu/person/ditzler
Bruce Fulton	iSchool, Teaches INFO 570 Database Development and Management	https://profiles.arizona.edu/person/bfulton
Gus Hahn-Powell	Linguistics, Teaches NLP	https://profiles.arizona.edu/person/hahnpowell
Mike Hammond	Linguistics, Teaches Speech Technology	https://profiles.arizona.edu/person/hammond
John Hartman	Computer Science, Teaches data science courses	https://profiles.arizona.edu/person/jhh

Bryan Heidorn	iSchool, Teaches data science courses	https://profiles.arizona.edu/person/heidorn
Peter Jansen	iSchool, Teaches NLP and other data science courses	https://profiles.arizona.edu/person/pajansen
Chris Lukeinbeal	Geography and Development, Teaches Geographic Info Systems	https://profiles.arizona.edu/person/clukinbe
Clayton Morrison	iSchool, Teaches machine learning and other data science courses	https://profiles.arizona.edu/person/claytonm
Jason Pacheco	Computer Science, Teaches data science courses	https://www2.cs.arizona.edu/~pachecoj/
Adriana Picoral	iSchool, Teaches data visualization and other data science courses, developing a new core course for the degree plan.	https://ischool.arizona.edu/sites/ischool.arizona.edu/files/Picoral_cv.pdf
Ryan Rucker	iSchool, Teaches INFO 570 Database Development and Management	https://ischool.arizona.edu/sites/ischool.arizona.edu/files/Ryan%20Rucker%20CV.pdf
Young-Jun Son	SIE, Professor and Department Head	https://profiles.arizona.edu/person/son
Meaghan Wetherell	iSchool, Teaches a variety of data science courses.	https://ischool.arizona.edu/sites/ischool.arizona.edu/files/CV%20_07022020%20_Meaghan%20Wetherell.pdf
Chicheng Zhang	Computer Science, Teaches a variety of data science courses.	https://zcc1307.github.io/

NOTE: as additional elective certificates are brought into the program this list will grow and will be submitted through the proper channels.

VII. SAMPLE PLAN OF STUDY- provide a sample plan of study.

Fall – Spring – Summer – Fall Plan (1.5 years)

Fall: INFO 520 Ethical Issues in Information and INFO 523 Data Mining and Discovery

Spring: INFO 526 Data Visualization, elective/certificate course #1, elective/certificate course #2

Summer: elective/certificate course #3 and elective/certificate course #4

Fall: elective/certificate course #5, elective/certificate course #6 and capstone project

Fall – Spring – Fall – Spring Plan (2 years)

Fall: INFO 520 Ethical Issues in Information and INFO 523 Data Mining and Discovery

Spring: INFO 526 Data Visualization, elective/certificate course #1, elective/certificate course #2

Fall: elective/certificate course #3 and elective/certificate course #4, elective/certificate course #5

Spring: elective/certificate course #6 and capstone project

VIII. STUDENT LEARNING OUTCOMES AND CURRICULUM MAP—describe what students should know, understand, and/or be able to do at the conclusion of this major. Work with Office of Instruction and Assessment to create a curricular map using Taskstream. Include your curricular map in this section (refer to Appendix C for sample Curriculum Map generated using Taskstream).

This M.S. in Data Science is intended to produce graduates who are knowledgeable in data science practices and have working skills that make them top candidates for data scientist positions. This degree will prepare students to participate in solving significant societal problems, whether within Arizona or globally, and to work effectively across disciplinary boundaries in a variety of careers in the public or private sectors.

Topics covered:

Dealing with Data

Types of Data

Data Quality and Preprocessing

Ethical Issues in Data Science

Managing, Storing, Preserving Data

Making Data Findable and Useable for People

Statistical Foundations of Data Science

- Probability
- Optimization
- Parametric and non-parametric statistical models

Machine Learning

- Classification
- Regression
- Clustering

Data Visualization

- Plots
- Maps
- Interactivity

Curriculum Map:

Shared as an individual document.

IX. ASSESSMENT PLAN FOR STUDENT LEARNING- using the table below, provide a schedule for program assessment of intended student learning outcomes 1) while students are in the program and 2) after completion of the degree.

Learning Outcomes	Sources(s) of Evidence	Assessment Measures	Data Collection Points
DS1: Students will demonstrate skills in processing and analyzing data.	Course-embedded assessments in INFO 523, the Capstone Project, and in several of the other course choices.	Exams, papers, created material, and other forms of student work	During each course, end of each course, and capstone team project.
DS2: Students will communicate with and effectively work and interact in teams.	Course-embedded assessments in INFO 523, the Capstone Project, and in several of the other course choices.	Group projects, oral and written communication assignments, final projects.	During each course, end of each course, and capstone team project.
DS3: Students will demonstrate abilities in analyzing ethical concerns and societal impacts related to data science.	Course-embedded assessments in INFO 520, and in several of the other course choices.	Exams, papers, and other forms of student work Summative critical self-reflections	During each course, end of each course, and capstone team project.

X. PROGRAM ASSESSMENT PLAN- using the table below, provide a schedule for program evaluation 1) while students are in the program and 2) after completion of the degree.

Assessment Measure	Source(s) of Evidence	Data Collection Point(s)
Job placement statistics	Student/alumni surveys	At graduation and as part of alumni survey
Academic program review	Reviewers' responses	Every 7 years
Student interest	Enrollment numbers	Every year
The School's academic success	National ranking	Every year

XI. ANTICIPATED STUDENT ENROLLMENT- complete the table below. What concrete evidence/data was used to arrive at the numbers?

5-YEAR PROJECTED ANNUAL ENROLLMENT					
	1st Year	2nd Year	3rd Year	4th Year	5th Year
Number of Students	50	100	200	400	500

Data/evidence used to determine projected enrollment numbers:

Projections are based on iSchool growth generally, a set of sample enrollment trends in similar degree programs across iSchools, and then we used an average of these numbers in our estimations.

XII. ANTICIPATED DEGREES AWARDED- complete the table below, beginning with the first year in which degrees will be awarded. How did you arrive at these numbers? Take into consideration departmental retention rates. Use National Center for Education Statistics College Navigator to find program completion information of peer institutions offering a same or similar program.

PROJECTED DEGREES AWARDED ANNUALLY					
	1st Year	2nd Year	3rd Year	4th Year	5th Year
Number of degrees	*n/a*	*45*	*90*	*180*	*360*

Data/evidence used to determine number of anticipated degrees awarded annually:

Projections are based on MS degrees, iSchool growth, and enrollment nationally. iSchool enrollment is among the fastest growing in many universities. We are expecting a high retention rate in the program, due to the high demand for and retention in our own current graduate level MS Information program and other similar data science-oriented programs across the campus.

General Demand

Students, generally, are living amid a massive shift in the amount of data we can save, use, analyze, and visualize – the Arizona region and students nationally thus need to be prepared for life and work in this data-driven economy:

- The data volumes are exploding; more data has been created in the past two years than in the entire previous history of the human race.

- Data is growing faster than ever before and by the year 2020, about 1.7 megabytes of new information will be created every second for every human being on the planet.

- By then, our accumulated digital universe of data will grow from 4.4 zettabyets today to around 44 zettabytes, or 44 *trillion* gigabytes.

- Every second we create new data. For example, we perform 40,000 search queries every second (on Google alone), which makes it 3.5 searches per day and 1.2 trillion searches per year.

- In Aug 2015, over 1 billion people used Facebook **FB +1.31%** in a single day.

- Facebook users send on average 31.25 million messages and view 2.77 million videos every minute.

- We are seeing a massive growth in video and photo data, where every minute up to 300 hours of video are uploaded to YouTube alone.

- In 2015, a staggering 1 trillion photos will be taken and billions of them will be shared online. By 2017, nearly 80% of photos will be taken on smart phones.

Source: http://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/#22f2f71c6c1d

Related Positions:

Artificial Intelligence Engineer
Business Analyst
Business Data Analyst
Business Intelligence Analyst
Business Intelligence Engineer
Data Associate
Data Analyst
Data Architect
Data Engineer
Data Scientist
Language Engineer
Machine Learning Engineer
Machine Learning Scientist
Market Research Analyst
Predictive Analytics Professional
Research Scientist

Local worksites for data-trained students include:

Apple Computer
Adobe
AdviNow Medical
Air Force Research Labs (Mesa AZ)
American Express
Databricks
IBM
Leidos
Lum AI
OpenClass
Pitch Vantage
State Farm

XIII. PROGRAM DEVELOPMENT TIMELINE- describe plans and timelines for 1) marketing the program and 2)student recruitment activities.

This program will be marketed alongside our other degree programs. As an iSchool we invest in event sponsorships so that we can hand out flyers and other marketing materials, we attend conferences, and advertise in print outlets and on the radio across Arizona. We plan to directly recruit students in and from locations such as:

- Existing undergraduate programs across the campus.

- UA events like the UA hackathon, or community events like TenWest.

- Social media

- The iSchool already collaborates with a number of cross-campus data science groups (e.g., the Data Science Institute, Cyverse) and aims to work collaboratively with those units for the purposes of advertising, recruitment, etc.

Upon approval, the School of Information will begin marketing and recruiting efforts immediately, accepting majors as soon as the program is approved.

XIV. DIVERSITY AND INCLUSION-describe how you will recruit diverse students and faculty to this program. In addition, describe retention efforts in place or being developed in order to retain students.

Student diversity in recruitment will be ensured through outreach activities that target college campuses that serve underrepresented student populations. The Curriculum and Instruction Committee will aim to increase diversity among the accepted students. Program information will be highlighted on the iSchool website, so that prospective students easily find it. The University of Arizona's diversity initiatives will be made visible on the website as well, with links that direct prospective students to these resources, so that they become aware of an existing support network for diversity and inclusion. Social media posts that aim to increase awareness about the proposed program will also encourage diversity. Outreach activities such as programming workshops have been held by iSchool faculty in the UAZ library, in the iSchool, and in other venues both on and off campus. These activities will be continued, as they help in increasing diversity and inclusion, in addition to outreach. We believe the current diverse student population of the iSchool will also encourage diverse student populations to apply. The 2019 racial composition of the iSchool was roughly 53% white, 19% Hispanic, 8% international, 7% Asian, 5% two or more races 5%, American Indian 1%, less than 1% unknown, less than 1% Pacific Islander. The iSchool's Knowledge River program, which aims to increase and maintain diversity will be another important factor in supporting and getting word out to underrepresented students. With all of these mentioned efforts, equitable access to the program will be ensured for a diverse and qualified pool of candidates, such as ethnic minorities and first generation and low-income students. Moreover, for the enrolled students, a nondiscriminatory and inclusive environment will always be maintained to provide support for students and increase their sense of belonging. To ensure an inclusive climate, diversity will also be emphasized in hiring of new faculty. Existing faculty will be encouraged to use inclusive materials in their courses and encourage their students to use inclusive materials in their coursework as well.

NEW ACADEMIC PROGRAM-GRADUATE MAJOR

XV. ABOR REQUIREMENT: New Academic Program Request. This section is required by ABOR. Most of the information can be copied/pasted from completed sections above. Instructions/clarification for completing the table below, from ABOR, can be viewed/downloaded here.

University: University of Arizona

Name of Proposed Academic Program: M.S. program in Data Science
Academic Department: School of Information
Geographic Site: Arizona Online
Instructional Modality: Online
Total Credit Hours: 30
Proposed Inception Term: Fall 2021

THE UNIVERSITY OF ARIZONA

NEW ACADEMIC PROGRAM-GRADUATE MAJOR

Brief Program Description:

Through transdisciplinary stacked certificate curriculum, the MS in Data Science program produces generations of professionals and researchers who have the knowledge, skill, and ability to engage in data science processes and work. This degree will encompass a blend of disciplines to address the grand challenges of today and the future. The MS in Data Science provides students the training they need in data collection, exploration, manipulation and storage, analysis, and presentation in order to navigate data-rich workplace environments.

Learning Outcomes and Assessment Plan:

DS1: Students will demonstrate skills in processing and analyzing data.

DS2: Students will communicate with and effectively work and interact in teams.

DS3: Students will demonstrate abilities in analyzing ethical concerns and societal impacts related to data science.

Student Learning Outcomes will be assessed annually through:

- Data-related projects, presentations, and visualizations produced relative to students' coursework.
- Regular survey of skills, abilities, and responsibilities of program graduates and employers of the graduates.

Projected Enrollment for the First Three Years:

5-YEAR PROJECTED ANNUAL ENROLLMENT					
	1st Year	2nd Year	3rd Year	4th Year	5th Year
Number of Students	50	100	200	400	500

Evidence of Market Demand:

The proposed MS in Data Science program will support the priorities outlined in the UA's Strategic Plan. Enhancing the supply of well-trained leaders in the data sciences will help catalyze economic development in the state of Arizona. Built upon the foundations of probability theory, optimization, as well as algorithmic thinking and doing, the data sciences include a broad set of approaches to include bioinformatics, natural language processing, machine learning, data mining, data visualization, and more.

According to U.S. News, Data Scientist jobs are 8 out of 100 best jobs in the United States. With a median salary of $94,280 and a 3.5% unemployment rate, this newer career field is increasingly recognized by employers as a valuable asset to their companies (*Data Scientist Ranks Among Best Jobs of 2021*, n.d.) The Bureau of Labor Statistics projects 31 percent employment growth for data scientists between 2019 and 2029 (*Fastest Growing Occupations: Occupational Outlook Handbook,* 2020).

Similar Programs Offered at Arizona Public Universities:

NAU has a Data Science graduate certificate (but not a full MS):
http://catalog.nau.edu/Catalog/details?plan=INFDSCT&catalogYear=2021

ASU has:

Master of Science in Business Data Analytics https://asuonline.asu.edu/online-degree-programs/graduate/master-science-business-analytics/

Master of Computer Science (Big Data Systems)

https://asuonline.asu.edu/online-degree-programs/graduate/master-big-data/

Master of Science in Program Evaluation and Data Analytics

https://asuonline.asu.edu/online-degree-programs/graduate/program-evaluation-and-data-analytics-ms/

New Resources Required? (i.e. faculty and administrative positions; infrastructure, etc.):

There are not external funds available for this M.S. program at this time, however a committee will be formed to provide focus on seeking external funds to support the School.

Program Fee/Differentiated Tuition Required? No

Specialized Accreditation? No

University of Arizona AMS
DEMO AREA

MS Data Science

Courses and Activities Mapped to MS Data Science

	Outcome		
	DS1 Students will demonstrate skills in processing and analyzing data.	DS2 Students will communicate with and effectively work and interact in teams.	DS3 Students will demonstrate abilities in analyzing ethical concerns and societal impacts related to data science.
Courses and Learning Activities			
INFO 520 Ethical Issues in Information Course assignments		IPA	IPA
INFO 521 Introduction to Machine Learning Course assignments	IPA		
INFO 523 Data Mining and Discovery Course assignments	IPA		IPA
INFO 526 Data Visualization (New Course) Course assignments	IPA	IPA	IPA
Capstone Course assignments	A	A	A
Survey Exit survey (Indirect)	A	A	A

Legend : I Introduced P Practiced A Assessed I/P Introduced/Prac

Last Modified: 02/08/2021 02:27:49 PM

NEW ACADEMIC PROGRAM-GRADUATE MAJOR

Graduate Major Peer Comparison Chart – select two peers for completing the comparison chart from (in order of priority) ABOR-approved institutions, AAU members, and/or other relevant institutions recognized in the field. The comparison chart will be used to identify typically required coursework, themes, and experiences for majors within the discipline. The comparison programs are not required to have the same degree type and/or major name as the proposed UA program. Information for the proposed UA program must be consistent throughout the proposal documents.

Program name	**Proposed UA Program:** M.S. Data Science	**Peer 1:** Master of Information and Data Science UC Berkeley School of Information	**Peer 2:** M.S. Data Science University of Wisconsin
Current # of enrolled students	[redacted]	792	393
Major Description.	The **MS in Data Science** will provide students the confidence and training they need in data collection, exploration, manipulation and storage, analysis, and presentation in order to navigate data-rich workplace environments. The degree will signal to employers that students have dedicated the time and energy necessary to develop the skills and confidence for tackling messy data problems using modern programming languages. The degree will service a diverse student population, training both 1) technically-minded students the nuances associated with successfully developing and communicating data methods and results, and 2)	The online Master of Information and Data Science (MIDS) is designed to educate data science leaders. The professional degree program prepares students to derive insights from real-world data sets, use the latest tools and analytical methods, and interpret and communicate their findings in ways that change minds and behaviors. The program features a multidisciplinary curriculum that draws on insights from the social sciences, computer	In the competitive world of data science, a master's degree is a requirement for advanced positions at top companies. The online UW Master of Science in Data Science is a smart choice for busy adults who want to advance their careers—or start a whole new career—but don't have time for on-campus courses. This data science master's program will teach you how to harness the power of big data using the latest tools and analytical methods. Courses are taught by diverse and

	less technically-minded students the basic skills necessary for gathering insights from data.	science, statistics, management, and law.	distinguished faculty from across the University of Wisconsin System. Tuition for the program is a flat fee of $850 per credit whether you live in Wisconsin or out of state, and financial aid is available for students who qualify. Because the program is entirely online, you can study and do homework whenever you like, wherever you have an Internet connection. Courses have no set meeting times and you never need to come to campus. An innovative Virtual Lab lets you remotely access software tools and programming languages such as R, Python, SQL Server, and Tableau, saving you the cost, time, and hassle of purchasing and installing these applications on your own computer.
Target careers	Artificial Intelligence Engineer Business Analyst Business Data Analyst Business Intelligence Analyst Business Intelligence Engineer Data Associate	Data scientist Data analyst Data engineer	Data scientist careers in health care, computer science, information technology, retail, marketing, manufacturing, transportation, communication, education, insurance, finance, science, security, law enforcement, and more.

	Data Analyst Data Architect Data Engineer Data Scientist Language Engineer Machine Learning Engineer Machine Learning Scientist Market Research Analyst Predictive Analytics Professional Research Scientist		
Total units required	30	27	36
Pre-admission expectations (i.e. academic training to be completed prior to admission)	Applicants are expected to have completed undergraduate coursework in statistics. Coursework in calculus preferred but not required. The application will have 2 explicit questions asking the applicant to list coursework or professional experience that meets these criteria. 1. What academic coursework and/or professional experience demonstrates your strong quantitative and analytical reasoning abilities? 2. What academic coursework and/or professional experience demonstrates your experience with math and programming, including data structures, analysis of	To be eligible for the online master's program, you must meet the following requirements: - A bachelor's degree and a GPA above 3.0. - A high level of quantitative ability. This should be demonstrated by at least one of the following qualifications: ● Work experience that demonstrates your quantitative abilities ● Academic coursework that demonstrates	You may be eligible for admission to this program if you have completed a bachelor's degree with a 3.0 or better grade point average (GPA) and prerequisite courses in elementary statistics, computer programming, and database administration.

	algorithms, and linear algebra?	quantitative aptitude - A working knowledge of fundamental concepts. For the MIDS program, this includes knowledge of data structures, algorithms and analysis of algorithms, and linear algebra. Applicants who lack this experience in their academic or work background but meet all other requirements for admission will be asked to complete a bridge course before enrolling in the Applied Machine Learning course. - Proficiency in programming languages, such as Python or Java, should be demonstrated by prior work experience or advanced coursework. Applicants who lack this experience in their academic or work background but meet all other admission requirements will be required to take the Introduction to Data Science Programming course in their first term.	

Major requirements **List all major requirements including core and electives.**	**Core - four courses** INFO 520 Ethical Issues in Information INFO 521 Introduction to Machine Learning INFO 523 Data Mining and Discovery INFO 526 Data Visualization (New Course) **Electives - choose five courses from this set** ECE 523 Machine Learning ECE 524 Fundamentals of Cloud Security ECE 579 Artificial Intelligence INFO 514/POL 514 Computational Social Science INFO 531 Data Warehousing and Analytics in the Cloud INFO 536 Data Science and Public Interests INFO 555 Applied Natural Language Processing INFO 556 Text Retrieval and Web Search INFO 557 Neural Networks INFO 570 Database Development and Management INFO 578/LIS 578 Science Information and its Presentation	**Foundation Courses:** Introduction to Data Science Programming (3 Units) Research Design for Application for Data and Analysis (3 Units) Statistics for Data Science (3 Units) Fundamentals of Data Engineering (3 Units) Applied Machine Learning (3 Units) Advanced Courses: Experiments and Causal Inference (3 Units) Behind the Data: Humans and Values (3 Units) Deep Learning in the Cloud and at the Edge (3 Units) Statistical Methods for Discrete Responses, Time Series, and Panel Data (3 Units) Machine Learning at Scale (3 Units)	DS 700: Foundations of Data Science DS 705: Statistical Methods DS 710: Programming for Data Science DS 715: Data Warehousing DS 730: Big Data: High-Performance Computing (Prerequisite: DS 710) DS 735: Communicating About Data DS 740: Data Mining & Machine Learning (Prerequisites: DS 705 and DS 710) DS 745: Visualization and Unstructured Data Analysis (Prerequisites: DS 700 and DS 740) DS 760: Ethics of Data Science (Prerequisites: DS 700 or DS 780) DS 775: Prescriptive Analytics (Prerequisite: DS 705) DS 780: Data Science and Strategic Decision Making DS 785: Capstone (Prerequisites: DS 715, DS 730, DS 735, DS 745, and DS 775)

	LING 508 Computational Techniques for Linguists LING/CSC/INFO 539 Statistical Natural Language Processing LING 582 Advanced Statistical Natural Language Processing SIE 530: Engineering Statistics SIE 533: Fundamentals of Data Science for Engineers SIE 545 Fundamentals of Optimization SIE 640 Large-Scale Optimization SIE 645 Nonlinear Optimization	Natural Language Processing with Deep Learning (3 Units) Data Visualization (3 Units) Capstone Course (3 Units)	

Research methods, data analysis, and methodology requirements (Yes/No). If yes, provide description.	Research course embedded in the set of core required courses.	Research course embedded in the set of core required courses.	Research course embedded in the set of core required courses.
Internship, practicum, applied course requirements (Yes/No). If yes, provide description.	Yes. Three capstone units are required. Students will complete a project that shows their mastery of data science skills.	Yes. MIDS students complete a capstone by executing a culminating project that integrates the core skills and concepts learned throughout the program. The capstone combines the technical, analytical, interpretive, and social dimensions required to design and execute a full data science project. Students learn integral skills that prepare them for long-term professional success in the field.	Yes. Capstone units required as part of the core curriculum. Capstone course in which students will develop and execute a project involving real-world data. Projects will include: formulation of a question to be answered by the data; collection, cleaning and processing of data; choosing and applying a suitable model and/or analytic method to the problem; and communicating the results to a non-technical audience.
Master thesis or dissertation required (Yes/No). If yes, provide description.	No	No	No
Additional requirements (provide description)			

THE UNIVERSITY OF ARIZONA®

BUDGET PROJECTION FORM

Name of Proposed Program or Unit: Data Science MS

	Projected		
Budget Contact Person: Amy Gordon	**1st Year** 2021 - 2022	**2nd Year** 2022 - 2023	**3rd Year** 2023 - 2024
METRICS			
Net increase in annual college enrollment UG			
Net increase in college SCH UG			
Net increase in annual college enrollment Grad	50	100	150
Net increase in college SCH Grad	600	1,200	1,800
Number of enrollments being charged a Program Fee	-	-	-
New Sponsored Activity (MTDC)	-		
Number of Faculty FTE	-	-	-
FUNDING SOURCES			
Continuing Sources			
UG RCM Revenue (net of cost allocation)			
Grad RCM Revenue (net of cost allocation)			
Program Fee RCM Revenue (net of cost allocation)			
F and A Revenues (net of cost allocations)			
UA Online Revenues	291,000	582,000	873,000
Distance Learning Revenues			
Reallocation from existing College funds (attach description)			
Other Items (attach description)			
Total Continuing	**$ 291,000**	**$ 582,000**	**$ 873,000**
One-time Sources			
College fund balances			
Institutional Strategic Investment			
Gift Funding			
Other Items (attach description)	500	250	250
Total One-time	**$ 500**	**$ 250**	**$ 250**
TOTAL SOURCES	**$ 291,500**	**$ 582,250**	**$ 873,250**
EXPENDITURE ITEMS			
Continuing Expenditures			
Faculty	86,000	172,000	258,000
Other Personnel			
Employee Related Expense	27,520	55,040	82,560
Graduate Assistantships			
Other Graduate Aid			
Operations (materials, supplies, phones, etc.)			
Additional Space Cost			
Other Items (attach description)			
Total Continuing	**$ 113,520**	**$ 227,040**	**$ 340,560**
One-time Expenditures			
Construction or Renovation			
Start-up Equipment			
Replace Equipment			
Library Resources			
Other Items (attach description)	500	250	250
Total One-time	**$ 500**	**$ 250**	**$ 250**
TOTAL EXPENDITURES	**$ 114,020**	**$ 227,290**	**$ 340,810**
Net Projected Fiscal Effect	**$ 177,480**	**$ 354,960**	**$ 532,440**

THE UNIVERSITY OF ARIZONA
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
School of Geography, Development & Environment

ENR2 Building, South 4th Floor
PO Box 210137
Tucson, Arizona 85721-0137
Ofc: 520-621-1652
Fax: 520-621-2889
geography.arizona.edu

April 11, 2021

Dear Curricular Review Committees,

This is a letter of support for the use of the School of Geography, Development and Environment's (SDGE), Geographic Information Systems and Technology (GIST) courses in the new Data Science M.S. (MS-DS) program.

SGDE supports the development of a MS-DS degree and our curriculum committee will be reviewing the best ways to participate using its current selection of GIST data science courses. SGDE can immediately participate in the success of the MS-DS program through the use of the Professional GIST (P-GIST) certificate which now allows any 3 graduate classes in GIST to be a certificate. We attach the list of courses for the P-GIST certificate.

Our curriculum committee is planning to review the use of the P-GIST and the possibility of creating a new geospatial data science (GDS) certificate in Fall 2021. The P-GIST or GDS certificate can be an option for the MS Data Science students and we hope the changes will be routed and reviewed in time for a spring 2022 launch. We note that a new data science faculty member is joining us in Fall 2021 – Beth Tellman who is an expert in the use of large data sets, including remote sensing, for earth system and flood analysis and was hired as part of the university-wide data science initiative.

We look forward to this and other collaborative efforts between our units.

Certificate 3: Database Management (School of Information and School of Geography, Development and Environment)
INFO 579: Database Design in SQL (under review/forthcoming)
INFO 570: Database Development and Management
GIST 570: Geodatabase (new course forthcoming)

Certificate 4: Geospatial Data Science (School of Geography, Development and Environment)
GIST 516e: Geovisualization
GIST 603A: Geographic Information Systems Programming and Automation
GIST 604B: Open Source GIS

Yours Sincerely,

Diana Liverman

Diana Liverman
Regents Professor
Director School of Geography, Development, and Environment

PGIST

P-GIST students can select 3 courses from the list below with the advance approval from a GIST advisor.

Course Descriptions

GIST 601A: Geographic Information Science (3 units)
This course will introduce the fundamental concepts of geographic information systems technology (GIST). It will emphasize equally GISystems and GIScience. Geographic information systems are a powerful set of tools for storing, retrieving, transforming and displaying spatial data from the real world for a particular set of purposes. In contrast, geographic information science is concerned with both the research on GIS and with GIS. As Longley et al. (2001, vii) note, "GIS is fundamentally an applications-led technology, yet science underpins successful applications." This course will combine an overview of the general principles of GIScience and how this relates to the nature and analytical use of spatial information within GIS software and technology. Students will apply the principles and science of GIST through a series of practical labs using ESRI's ArcGIS software.

GIST 601B: Remote Sensing Science (3 units)
This course provides an introduction to the scientific principles and practices of remote sensing. Topics that will be covered in this course include issues of spatial resolutions, the electromagnetic spectrum, remotely sensed sensors, spectral characteristics, digital and digitalization issues, multispectral and LiDAR image processing and enhancement, and land-use and land-cover classifications (LULC) and change detection. The course also emphasizes integration issues and analysis techniques that arise when merging remotely sensed data with geographic information systems (GIS).

GIST 602A: Raster Spatial Analysis (3 units)
This course examines the principles and practices associated with raster data development and analysis, particularly the development of real world surfaces and statistical analysis based on these surfaces. The course is presented in a lecture/laboratory format. The lecture portion will deal with conceptual issues necessary for the use of raster approaches within a GIS framework. The laboratory portion will provide practical experience with rasters in an ArcGIS environment.

GIST 602B: Vector Spatial Analysis (3 units)
This course focuses on providing students with an introduction vector based spatial analysis and their application in GIS software. Students will learn about how to analyze distribution, direction, orientation, clustering, spatial relationships and processes, and how to render analytic outcomes into cartographic form. This course provides foundational knowledge of global positioning systems, data collection, geodatabase development, and georeferencing.

GIST 603A: Geographic Information Systems Programming and Automation (3 units)
The goal of this course is to gain an introductory understanding of geographic programming and data automation techniques using ModelBuilder and the Python language. Students will become familiar with the ModelBuilder tools inside ArcGIS for Desktop to automate redundant tasks using ModelBuilder and learn how to build a script using Python to customize functionality and task with GIS.

GIST 603B: WebGIS (3 units)
The goal of this course is to gain an understanding of web mapping using applications like ArcGIS for Server, ArcGIS Online (AGOL), WebAppBuilder (WAB), web-enabled geoprocessing, Story Maps, AppStudio, and the Javascript API.

GIST 604A: Applied GIS (3 units)
A GIST-based problem solving approach within the context of a student-directed project. Specific GIS skills covered include project planning, spatial data sources and acquisition, data compilation, coding, analysis, representation, and presentation of results. The course can be repeated for credit, as the topics will vary; each course will examine a different urban or environmental issue in the natural and social sciences using geographic information systems technology.

GIST 604B: Open Source GIS (3 units)
This course provides students a brief introduction about Open Source software for both desktop and internet GIS applications. Main objective of the course is to expose students to alternative open source tools for practicing GIS besides licensed and conventional GIS software. Students will go through hands on learning about applications hosting, data development, processing, and sharing using open source tools and technologies such as GITHub , Quantum GIS (QGIS), Python, GeoServer and PostGIS. Students will apply technology in lab assignments using real-world data.

THE UNIVERSITY OF ARIZONA

Systems & Industrial
Engineering Department

1127 E James E Rogers Way
P.O. Box 210020
Tucson, AZ 85721-0020
(520) 621-6551
Fax: (520) 621-6555
www.sie.arizona.edu

January 25, 2021

Dear Curricular Review Committees,

This is a letter of support for the use of five of our Systems and Industrial Engineering (SIE) courses to begin as part of the new online Data Science M.S. program. We see the need for this new program at the University of Arizona, and it is a program we are happy to be a part of. In the near future we would like a chance to re-examine the curricular needs in the program so that we can refine the course plan based on further discussions among our faculty in those areas and also given the enrollment realities we face in certain courses.

There are many great opportunities for synergy between Systems and Industrial Engineering and the iSchool moving forward. So, we are delighted to collaborate across units for this effort.

Sincerely,

Young-Jun Son
Professor and Head of Department of Systems and Industrial Engineering

ELECTRICAL & COMPUTER ENGINEERING
College of Engineering
1230 E. Speedway Blvd.
P.O. Box 210104
Tucson, AZ 85721-0104

Ofc: 520.621.6193

ece.engineering.arizona.edu

COLLEGE OF ENGINEERING
Electrical & Computer Engineering

January 24, 2021

Dear Curricular Review Committees:

This is a letter of support for the use of our courses (ECE 523 Machine Learning, ECE 524 Fundamentals of Cloud Security, and ECE 579 Artificial Intelligence) in the new online Data Science M.S. program. We see the need for this new program at the University of Arizona, and it is a program we are happy to be a part of. There are many great opportunities for synergy between Electrical and Computer Engineering and the iSchool moving forward. So, we look forward to this and other collaborative efforts between our units.

Sincerely,

Tamal Bose
Professor and Department Head

THE UNIVERSITY OF ARIZONA

DEPARTMENT
LINGUISTICS

College of Social & Behavioral Science
1103 E. University Blvd
PO Box 210025
Tucson AZ 85721-0025

Tel: 520-621-6897
Fax: 520-626-9017

http://linguistics.arizona.edu

January 22, 2021

Catherine Brooks
Director, School of Information
University of Arizona

Dear Dr. Brooks (Dear Catherine):

I am writing to state the Department of Linguistics' support for the proposed new online M.S. in Data Science program. We have discussed with you which Linguistics courses to include in the curriculum of the program, and we are in agreement. The proposal has the support of the Linguistics Curriculum Committee and the faculty who teach in this area. We look forward to collaborative work between this program and our own online M.S. in Human Language Technologies program.

Sincerely,

Natasha Warner
Professor and Department Head
Department of Linguistics

From: Sandoval, Liz - (esandmar)
To: Sandoval, Liz - (esandmar)
Subject: FW: Master"s degree in Data Science
Date: Tuesday, April 13, 2021 11:18:24 AM

From: Carnie, Andrew H - (carnie) <carnie@arizona.edu>
Sent: Tuesday, April 13, 2021 11:13 AM
To: Sandoval, Liz - (esandmar) <esandmar@arizona.edu>
Subject: Re: Master's degree in Data Science

Hi Liz,

Best,

Andrew

Begin forwarded message:

From: "Cheu, Elliott C - (echeu)" <echeu@arizona.edu>
Subject: Master's degree in Data Science
Date: April 13, 2021 at 8:51:48 AM MST
To: "Carnie, Andrew H - (carnie)" <carnie@arizona.edu>

Hi Andrew,

The College of Science supports the creation of a new Master's Degree in Data Science offered by the School of Information.

Best regards,

Elliott Cheu, Ph.D.
Interim Dean, College of Science
Distinguished Professor of Physics
University of Arizona
(520) 621-4092