

Table of Contents

General Information	2
Relationship to Institutional Mission and Statewide Goals	5
Program Description	7
Course of Study.....	12
Expected Learning Outcomes For Students And Means Of Assessment	15
Accreditation	17
Need for this Credential	18
Other Similar Programs	19
Program Integration And Collaboration	20
Financial Sustainability.....	22
Untitled Section – Resources	24
Other Program Characteristics.....	27

General Information

Give a brief (1-2 paragraphs) overview of the proposed credential, including its disciplinary foundations and connections, its focus and learning objectives for students, and the specific degree (e.g. bachelors, masters, doctorate) and/or credentials (e.g. major, certificate, minor, concentrations) to be offered. This should be based largely on your descriptions in the following sections, but it should be shorter than their combined length. Moreover, it should use language that is capable of communicating your ideas to audiences increasingly distant from your academic field as your proposal moves through the review process.

We propose an undergraduate degree in data science that is quantitatively focused, requiring courses in computer science, mathematics, probability and statistics, and machine learning. This quantitative core of courses integrates with applications of these concepts and tools in specific areas such as biology, finance, geography, linguistics, etc. (so called “domain” areas). The degree finishes with the option of a capstone project demonstrating mastery of data science concepts within the student's chosen domain of specialization.

Following examination of successful programs at other top universities, we closely followed the curriculum in the Division of Data Science and Information at UC Berkeley. This program is widely recognized as a model and has explicitly developed data science teaching material to be utilized by other programs and universities. We have adopted and modified proposed entry and mid-level courses in core data science proficiencies and exposure to domain emphases, as well as steps through key offerings in math, computer science, ethics, and communication. This curriculum will require the development of six new courses and the reorientation of a small number of relevant courses in math and CIS. These departments have agreed to be key partners in this major, and the Data Science Initiative has resources to support the development of additional classes.

The UO Data Science Visioning Committee (DSCV) concluded that a strong domain emphasis through collaboration with existing schools, colleges, and departments would be optimal for an undergraduate degree in data science. In accordance with this vision, we have developed a university-wide program that delivers a quantitative core of data science competency and enables collaboration through the domain-emphasis aspect of the degree. To be clear, the proposal here is to form a degree in data science that contains emphases in partnered domain areas. We are not proposing data science tracks in other units.

The core data science training is aligned with existing courses applying data science techniques in several units at UO. Students who wish to obtain a data science degree with an emphasis in a certain domain will be able to take those courses. The students will be required to take any prerequisites necessary for those courses that are required within that unit. This is common practice within other data science programs we examined, and the

students will be made aware of that fact. We are actively working with partner units to optimize prerequisites to enable data science students to take domain emphasis courses.

Because of increasing quantitative and computational approaches in many units and the more recent set of data science related hires, we have the opportunity to build key connections that allow the Data Science program to be both a self-contained data science degree with domain emphases, as well as a potential gateway for future data science tracks, minors, or majors in other units. This action is the purview of those departments, schools and colleges, and would require all necessary approvals.

To be clear, we are not proposing new degrees or courses of study in other units. Our focus for AY19/20 will be on creating the Data Science program and degree. In addition, we will be developing the processes by which additional domain emphases within the data science degree will be added and working with interested units to add those emphases. We are committed to working with units in future years to develop data science tracks within their degrees. Doing so might involve changes in prerequisites within those departments, as well as honing of course offerings within the Data Science program. This course of program development will allow the creation of a unique ecosystem at the University of Oregon: a network of domain emphases in a core data science degree, as well as data science tracks in a plethora of other degrees across the institution.

Primary Proposer

Joseph Sventek

Is there a co-proposer for this proposal?

Yes

Co-proposer(s)

Name	Home Unit
William Cresko	Data Science

Home department

Data Science

College

Arts & Sciences, College of

Additional Department Affiliations

Department
Computer & Information Science
Mathematics

Level

Undergraduate

Program Type

Degree

Degree Type

Bachelor

By default, the program will be approved for the Bachelor of Arts and Bachelor of Science. If you are only requesting one of these, please indicate below:

Primary Location

UO main campus

Program Delivery Format

Traditional classroom and dry laboratory

Does the program represent a collaboration of two or more university academic units?

Yes

Relationship to Institutional Mission and Statewide Goals

How is the program connected with the UO's mission, signature strengths and strategic priorities?

The University of Oregon is a comprehensive public research university committed to exceptional teaching, discovery, and service. Because data science is a growing interdisciplinary field, a data science undergraduate degree program is essential to UO's mission and is part of a key presidential initiative. The collaborative nature of this degree builds upon UO strengths in interdisciplinary programs.

How will the proposal contribute to meeting UO and statewide goals for student access and diversity, quality learning, research, knowledge creation and innovation, and economic and cultural support of Oregon and its communities?

Accessibility, equity and inclusion are fundamental to the data science program. By nature, data science involves and impacts all members of society. Correspondingly, an essential measure of success for the program is the recruitment and sustained support of a richly diverse student cohort, particularly for students traditionally underrepresented in STEM fields. We plan to intentionally include advising and activities to attract and support such a diverse community of students and scholars.

Specific efforts with respect to education and quality learning involve adopting a pedagogical approach that embraces students who may not otherwise be exposed to STEM fields or who may be intimidated by quantitative STEM programs. The first foundational data science course, DSCI 101, has been developed to minimize barriers to entry. It has no prerequisites and embraces students with no prior computing or statistical background. It will be offered as a UO core education course with the intent to have as broad of a reach and impact as possible. In-class demonstrations and examples are intentionally sourced to include culturally inclusive content and to address a variety of learning styles. As students progress through the lower- and upper-division DSCI curriculum, ethics and human context are woven throughout each of the DSCI courses.

Further initiatives will emulate the successful practices employed by the CIS department, including active recruitment of women and minorities into the program and summer internship placements with companies partnering with the data science program. It is expected that these measures will ensure successful employment of data science students, including those traditionally underrepresented, as has been the case for CIS.

How will the proposal meet regional or statewide needs and enhance the state' capacity to:

- **improve educational attainment in the region;**
- **respond effectively to social, economic and environmental challenges and opportunities;**
- and
- **address civic and cultural demands of citizenship?**

Experienced data scientists are a growing economic need. The January 2019 report from Indeed, one of the top job websites, showed a 29% increase in demand for data scientists year over year, and a 344% increase since 2013. Data from the technology job website Dice showed a 32% increase in data science job postings year over year. Dice also noted that the job postings are from companies in a wide variety of industries, not just technology – e.g., investment banking, insurance, healthcare. Indeed currently lists 89 open data science jobs in Oregon; if the current year-on-year growth rate continues, this means that there will be 317 open data science jobs on this date in 2024 when we produce our first group of graduates.

With so many different domains requiring data science expertise, the core + domain emphasis built into the proposed program will help students to develop the core quantitative skills and apply them immediately to their domain of interest. We expect graduates of this program to find quality jobs immediately after graduation, both in the local area and throughout Oregon and the rest of the US.

In alignment with the deep history of UO as a liberal arts research university, a goal of the program is to not only help students utilize data science but to think critically about the impacts of data science on society. The program requires one course in data ethics, with other ethics courses available as area electives; three data science core courses (DSCI 101, 102, 311) also include ethical topics in their syllabi; finally, courses in cybersecurity are also available as area electives. This is an important aspect of the program addressing civic and cultural demands of citizenship - securing and ethically using domain data is essential to prevent misuse of said data.

Program Description

Is there a core set of required courses?

Yes

What is the core set of required courses and what is the rationale for giving these courses this prominent role? What are the central concepts and/or skills you expect students to take from the core?

DSCI 101-102 - broad, general introduction to data science

CIS 210-212 - basic knowledge of computational thinking, object-oriented programming, and data structures

MATH 251-252 - sound mathematical knowledge of calculus

MATH 341-342 - sound mathematical knowledge of linear algebra

DSCI 311 - deep dive in particular areas of data science

DSCI/MATH 345 - probability and statistics for data science

PHIL 423 - technology ethics

DSCI/CIS 372 - machine learning for data science

Armed with the knowledge from the courses above, a student is in position to apply data science techniques in a multitude of domains.

What is the relationship between upper-division courses and the lower-division curriculum? For example, are fundamental principles introduced in the lower division and then applied to increasingly complex problems at the upper-division? This vertical architecture is common in the sciences but is by no means universal. In the humanities, a more horizontal structure is often appropriate. For example, students might read and analyze literature at each level (100-400) but do so with increasing sophistication and the capacity to draw on a widening array of literary forms and ideas.

Upper-division data science courses build upon lower-division courses. Within the domain emphases, introductory domain courses are first taken as part of the “domain core”; data-science-focused “domain specialization” courses are taken in the upper-division, when the student will have completed most of the core courses above. The final year capstone project is the culmination of the structured curriculum, with the student demonstrating the ability to apply all of the concepts and techniques acquired to a comprehensive data science project in the domain emphasis. The flow of lower division to upper division courses allows a core set of competencies and proficiencies in data science to be delivered to a wide student body that can then be applied in a diversity of domain emphases across the university.

Are there specific course-to-course prerequisites that help students extend or link ideas or are the intellectual connections among courses in your major more general?

CIS 210-211-212 must be taken in sequence.

MATH 251 and 252 must be taken in sequence. MATH 252 must be completed before starting MATH 341 and 342, which must be taken in sequence.

MATH 342 and CIS 211 must be completed before taking DSCI/MATH 345.

DSCI 101 and 102 must be taken in sequence. DSCI 102 must be completed before taking DSCI 311.

DSCI/CIS 372 must be taken after the completion of CIS 212, MATH 342, and DSCI/MATH 345.

DSCI 411, if taken, must be preceded by completion of DSCI/CIS 372 and PHIL 423.

Are there tracks or concentrations within the credential? If so, do these start from a common core or are they differentiated from the beginning?

An essential aspect of the degree in data science is that data science majors develop critical competencies in a domain emphasis of their choosing.

The domain emphasis consists of completing 2-3 courses (8-12 credits) in the domain core, followed by a minimum of 3 courses (12 credits) of domain specialization. For each domain emphasis, a curated list of courses has been developed for both the core and specialization component.

Currently, domain emphases have been established for biology, geography, accounting analytics, marketing analytics, and linguistics. The curated list of domain core and domain specialization courses for each domain is outlined below. The domain emphases have been developed in partnership between the Data Science program and the corresponding units in each domain (please see attached documentation).

It is important to note that some of these courses have their own set of prerequisites. Where possible, the Data Science program is working with corresponding units to optimize prerequisites to facilitate the process of data science majors completing domain-satisfying courses. The Data Science program will also provide its students with mappings of required prerequisites for the courses in each domain. Ultimately, students will be responsible for ensuring that they are able to complete the prerequisite courses required for the domain emphasis they select.

A domain-owning unit that is a part of the Data Science program agrees to the following:

- the unit works with the DS program to define the list of domain core courses
- the unit works with the DS program to define the list of (at least 4) domain elective courses, documenting constraints on course choice in this list

- DSCI students will have the same registration priority and access to courses in these lists as the unit's own majors
- the unit will expand offering of DSCI listed courses if these courses cannot accommodate new and existing students
- the unit participates in an annual update to the core and elective lists, usually in Winter quarter, for students to consider when enrolling for their courses in the subsequent academic year
- the unit will work with the DS program to optimize prerequisites for the domain elective courses
- the unit will declare if it supports joint supervision of students taking the Capstone course
- if the unit **does** support joint supervision of students taking the Capstone course:
 - the unit will identify faculty who will provide supervision for the Capstone course
 - the unit and the identified faculty will provide raw domain data for use by Capstone students

Biology Domain:

Domain Core – select at least two

- BI 211 – General Biology I – Cells 4 cr
- BI 212 – General Biology II – Organisms 4 cr
- BI 213 – General Biology III – Populations 4 cr

Domain Specialization – select at least three

- BI 360 – Neurobiology 4 cr
- BI 485 – Techniques in Computational Neuroscience 4 cr
- BI 320 – Molecular Genetics 4 cr
- BI 493 – Genomic Approaches and Analysis 4 cr
- BI 370 – Ecology 5 cr
- BI 471 – Population Ecology 4 cr

Geography Domain:

Domain Core – take all three

- GEOG 181 – Our Digital Earth 4 cr
- GEOG 281 – The World & Big Data 4cr
- GEOG 481 – GIScience I 4 cr (Formerly 200-level)

Domain Specialization – select at least three

- GEOG 482 – GIScience II 4 cr
- GEOG 485 – Remote Sensing I 4 cr
- GEOG 486 – Remote Sensing II 4 cr
- GEOG 490 – Special Topics (must have approval from the DSCI director of undergraduate studies) 4 cr
- GEOG 491 – Advanced GIS: Python 4 cr

- GEOG 493 – Advanced Cartography 4 cr
- GEOG 496 – Location-Aware Systems 4 cr
- GEOG 498 – Geospatial Project Design 4 cr

Accounting Analytics Domain

Domain Core – take all three

- BA 101 Introduction to Business 4 cr
- BA 215 Accounting: Language of Business Decisions 4 cr
- EC 201 Introduction to Economic Analysis: Microeconomics 4 cr

Domain Specialization (assumes a 4th course fulfills the DS capstone requirement, the ACTG 410 capstone below is designed to have a major capstone assignment)

Required:

- BA 240¹
- ACTG 350² Intermediate Accounting I 4 cr

Take two of these three:

- ACTG 410² Accounting Data & Analytics 4 cr
- ACTG 410 Accounting Data & Analytics Capstone 4 cr
- OBA 410³ Predictive Analytics 4 cr

Marketing Analytics Domain

Domain Core – take all three

- BA 101 Introduction to Business 4 cr
- BA 215 Accounting: Language of Business Decisions 4 cr
- EC 201 Introduction to Economic Analysis: Microeconomics 4 cr

Domain Specialization (assumes a 4th course fulfills the DS capstone requirement)

Required:

- BA 240¹
- BA 317 Marketing: Value for Customers 4 cr
- MKTG 390⁴ Marketing Research 4 cr

Take one of the following two:

- MKTG 395 Marketing Analytics 4 cr
- OBA 410³ Sports Analytics 4 cr

Linguistics Domain:

Domain Core – take both

¹ Students who already have significant experience with Excel can test out of BA 240.

² LCB is prepared to drop EC 202 as a pre-requisite, and alter the pre-requisites to accept “ACTG 213 or BA 215” and “MATH 243 or DSCI/MATH 345”

³ LCB is prepared to alter the pre-requisite to accept “OBA 312 or DSCI/MATH 345” and “OBA 335 or DSCI 311”

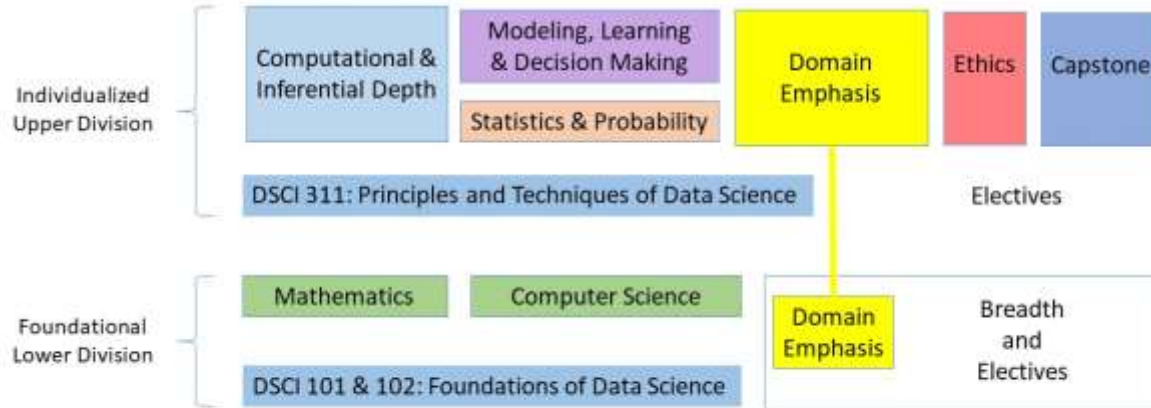
⁴ BA 317 is an allowable substitute for MKTG 311

- LING 301 – Linguistic Analysis 4 cr
- LING 302 – Linguistic Behavior 4 cr

Domain Specialization – select at least three

- LING 435 – Morphology & Syntax 4cr
- LING 451 – Functional Syntax I 4cr
- LING 452 – Functional Syntax II 4cr
- LING 493 – Corpus Linguistics 4cr
- CIS 410nlp – Natural Language Processing 4cr

General Course Outline



The above figure shows the general course outline for a student pursuing a degree in Data Science with a domain emphasis. Each of the boxes are described in additional detail in the following sections. An additional domain specialization course may be taken in lieu of the capstone senior project.

Course of Study

BS Flight Path

	Fall	Winter	Spring
Year 1	CIS 210	CIS 211 DSCI 101	CIS 212 DSCI 102 MATH 251
Year 2	Dom ^a 1 MATH 252	Dom 2 MATH 341	[Dom 3] MATH 342
Year 3	DSCI 311 DSCI/MATH 345	DSCI/CIS 372 PHIL 423	CID ^b 1
Year 4	Dom 4 CID 2	Dom 5 CID 3	Dom 6 DSCI 411 (or Dom 7)

^aDom = domain emphasis

^bCID = Computational and Inferential Depth

N.B. an additional course from the curated domain specialization list may be taken in place of the capstone senior project.

Course of Study

The data science curriculum combines general principles with domain-specific application. The curriculum is sub-divided into the following categories with the corresponding requirements:

Core data science courses:

- DSCI 101 – Foundations of Data Science I
- DSCI 102 – Foundations of Data Science II
- DSCI 311 – Principles and Techniques of Data Science
- DSCI 411 – Data Science Capstone Project*[‡]

*An additional domain-specialization course may be substituted in lieu of DSCI 411.

[‡] DSCI 411 will be restricted to data science majors with an average GPA of at least 3.75 over all data science degree courses.

Foundations in Mathematics and Computing:

- Foundational computer science series: CIS 210, 211, 212
- Foundational Calculus series: MATH 251, 252
- Linear Algebra Series: MATH 341-342

Computational and Inferential Depth

Students must select **three** courses from the following list below. Please see the note at the end of this section regarding CIS prerequisites in the Computational and Inferential Depth category.

- CIS 313 – Intermediate Data Structures
- CIS 314 – Computer Organization
- CIS 322 – Introduction to Software Engineering
- CIS 333 – Applied Cryptography
- CIS 451 – Intro to Databases
- CIS 315 – Intermediate Algorithms
- CIS 330 – C/C++ & Unix
- CIS 410vis – Visualization
- CIS 415 – Operating Systems
- CIS 432 – Intro to Internet
- CIS 434 – Distributed Systems and Network Security
- CIS 425 – Programming Languages
- CIS 422 – Software Engineering
- CIS 471 – Intro to AI

Modeling, Learning and Decision Making

- DSCI/CIS 372 – Machine Learning for Data Science

Probability

- DSCI/MATH 345 – Probability and Statistics for Data Science

Human Contexts and Ethics

- PHIL 423 – Technology Ethics

Domain Emphasis

The domain emphasis consists of completing 2-3 courses (8-12 credits) in the domain core, followed by a minimum of 3 courses (12 credits) of domain specialization. For each domain emphasis, a curated list of courses has been developed for both the core and specialization component. Please see the previous section (tracks/concentrations) for a detailed list of courses that satisfy each available domain emphasis.

Note on CIS Prerequisites in the Computational and Inferential Depth category:

The courses listed in the Computational and Inferential Depth category have the following prerequisites:

1. CIS 313, 314, 322, 333, and 451 have CIS 212 as a prerequisite; 313 has MATH 232 as a prerequisite, but DSCI students will have completed MATH 252 and 342, which provides evidence of the same mathematical sophistication; note that a proposal for CIS 333 has

been submitted to CASCC for consideration in Fall 2019. Thus, a student can choose 3 courses from this subset at the current prerequisite settings.

2. CIS 315 has CIS 313 as a prerequisite. Choosing to take 315 automatically locks in 313, so the student only has to take one more course from the list.
3. CIS 330 has CIS 314 as a prerequisite. Choosing to take 330 automatically locks in 314, so the student only has to take one more course from the list.
4. CIS 410viz has CIS 330 as a prerequisite. Choosing to take 410viz automatically locks in 330 and 314.
5. CIS 415 has CIS 330 and CIS 313 as a prerequisite. CIS will be proposing a change to the prerequisites for CIS 415 for consideration in Winter 2020 – both CIS 313 and 330 will be removed as prerequisites, and CIS 212 will be the lone prerequisite for CIS 415.

The remaining CIS courses that are listed have complicated prerequisites; CIS will be discussing these in this academic year, with an eye toward optimizing the prerequisites, but if a student were to choose a course out of this subset before that happens, they would have to satisfy the currently defined prerequisites.

Expected Learning Outcomes For Students And Means Of Assessment

Learning Outcome	Introduced	Developed	Mastery
1. Have demonstrated the ability to assess data set quality, identifying and rectifying potential errors in such a way so as to lead to statistically meaningful derived information	DSCI 101	DSCI 102	DSCI 311 DSCI 411
2. Be able to visualize complex datasets using descriptive statistics and graphical representations	DSCI 101	DSCI 102	DSCI 311 DSCI 411
3. Demonstrate understanding of basic regression, optimization, prediction, simulation, and visualization methods	DSCI 102	DSCI 311	DSCI 345
4. Be able to use critical thinking skills to translate substantive questions into well-defined statistical or probability problems and choose the appropriate graphical or numerical descriptive and/or inferential statistical techniques for a given problem, leading to actionable, valid, and meaningful conclusions	DSCI 311	DSCI 345	DSCI 411
5. Have developed successful strategies for formulating and testing hypotheses about data	DSCI 311	DSCI 345	DSCI 411
6. Have demonstrated an understanding of ethical, legal, societal, and economic concerns	DSCI 101 DSCI 102	DSCI 311	PHIL 423 DSCI 411
7. Be able to apply fundamental concepts of data science (data management, statistical prediction and inference, experimental design, etc.) to applications specific to the chosen specialization domain			DSCI 411

If needed, explain particular items in the grid:

Not applicable

If needed, describe your curriculum map in narrative form, as an alternate to the grid:

Not applicable

What is the nature and level of research and/or scholarly work expected of program faculty which will be indicators of success in those areas?

Tenure-track faculty appointments will be made in various units throughout the university so as to maximize the interdisciplinary research strengths of the combined faculty. Such faculty will be expected to conduct innovative and impactful research, demonstrated by a history of successfully awarded grants, journal and or conference publications, and invited presentations.

Career NTTF appointments will be made in the Data Science program, itself. There may also be TTF appointments in the program in the future if the research focus is on core data science research.

TTF and NTTF instructors are expected to deliver high-quality, engaging courses, as measured by student course evaluations, student success, and peer teaching evaluations.

Explain the methods by which the learning outcomes will be assessed and used to improve curriculum and instruction.

The assessment of degree learning outcomes will be derived from a sub-sample of artefacts generated through the program. The following will enable us to track each outcome from cohort to cohort:

1. Mastery of this learning outcome is covered in DSCI 311; one project that students must submit for assessment in 311 is to take a dirty data set, representing all the types of imperfections covered in the courses, and using techniques that they have learned in 101, 102, and 311, they must clean the data set for a standard statistical analysis.
2. One project that students must submit for assessment in 102 is to compute standard statistical attributes for a given data set and to use an appropriate graphical representation to communicate those attributes.
3. This will be assessed through one of the assessment mechanisms in DSCI 345.
4. This will be assessed through a detailed analysis of the DSCI 411 project reports, particularly establishing an hypothesis of information buried in the data, applying suitable and appropriate analysis techniques to the data, establishing the statistical significance of the results obtained, and suitably visualizing the results of the project for consumption by relevant stakeholders.
5. This will be assessed through one of the assessment mechanisms in DSCI 345.
6. This will be assessed through one of the assessment mechanisms in PHIL 423.
7. This will be assessed through a detailed analysis of the DSCI 411 project reports, particularly the usefulness of the inferred information to the domain.

A full assessment plan will be devised to schedule these assessments over a 3-year cycle. The anticipated schedule will look as follows:

Outcome	Year 1	Year 2	Year 3
1	x		
2		x	
3			x
4	x	x	x
5	x		
6		x	
7			x

Accreditation

Is or will the program be accredited?

No.

Please explain why accreditation is not being sought:

Data science is a new discipline and as such there is no accreditation standard or organization that accredits undergraduate data science programs. Once such an accreditation organization and standard has been developed, we will reconsider accrediting the program.

Need for this Credential

What are the expected degrees/certificates over the next five years.

Number of Degrees:

Year 1	Year 2	Year 3	Year 4	Year 5
0	0	0	25	100

What are possible career paths for students who earn this credential? Estimate the prospects for success of graduates in terms of employment, graduate work, licensure, or other professional attainments, as appropriate.

Data science is a growing field with increasing employment demands. Graduates of the UO's data science program will be competitive for data scientist positions in industry, including data analyst, quantitative analyst, data engineer, or artificial intelligence/machine learning engineer. The demand for these positions has grown substantially in all sectors with the increasing availability of business and customer data. The workforce analysis company Glassdoor estimates starting salaries between \$81k-\$149k for entry-level data scientists. UO's program requires experience in a domain area which would offer additional employment opportunities. A domain emphasis in biology, for example, would prepare graduates for careers in bioinformatics or biostatistics. Qualified students will have the option to complete a capstone project to demonstrate competency to potential employers, which would increase their competitiveness in the job market.

The level of computer science and mathematics study in the data science program is most likely insufficient to win a PhD place in either of those two core disciplines. It is also unlikely that there will be many data science PhD programs, and therefore students pursuing a data science undergraduate degree are unlikely to enter a postgraduate data science PhD degree program. Students majoring in Data Science wishing to pursue a PhD in a core discipline are advised to double major in their PhD target discipline. However, because of the increasingly quantitative nature of nearly all fields of advanced study, students who pursue a data science undergraduate degree with a strong domain application (e.g. biology) in fulfillment of the domain emphasis are likely to be well placed to enter PhD programs in that domain area.

Other Similar Programs

Are there similar or related programs currently offered at the University of Oregon?

No

Attach your communications showing due diligence in consulting with other UO departments or areas

See attached file, "Due Diligence – Combined.pdf"

List any existing programs that are complemented or enhanced by the new major.

CIS offers experimental courses in data science at a core education level and at an upper division level. These courses will be removed from the CIS offerings once this program is operational.

Describe the steps that have been taken to ensure that the proposed program(s) does not overlap other existing UO program(s) or compete for the same population of students. [Provide documentation that relevant departments or areas have been informed of the proposal and have voiced no objections.]

The only units that have any overlap are:

- Mathematics – this department is a partner in the delivery of the proposed program
- Computer and Information Science – this department is a partner in the delivery of the proposed program. The CIS department is working to optimize prerequisite requirements on its courses that are relevant to this degree.
- Biology, Geography, Linguistics, Accounting Analytics, Marketing Analytics – these units are the initial partners in the delivery of the domain emphasis aspect of the proposed program

Please see the attached documentation which shows due diligence of the Data Science program in coordinating with the above units

Program Integration And Collaboration

Are there closely-related programs in other Oregon public or private universities?

Yes

List similar programs and indicate how the proposal complements them. Identify the potential for new collaboration.

Of the four major research universities in the state of Oregon (UO, OHSU, PSU and OSU), only OSU is proposing an undergraduate program with a domain emphasis structure that may be comparable to UO's. The OSU program in biological data science is a natural outgrowth of the Computational Genomics and Research Biotechnology (CGRB) program that has existed at OSU for many years, and with which there have been numerous connections with UO. The new OSU program focuses on biological data science and offers three options, including: computational biology, ecological and environmental informatics, and genomics. Each of these options are focused in the natural sciences. In contrast, the UO program currently being proposed has been designed to be inclusive of disciplines across the UO and currently includes domain emphases in the natural sciences (biology), social sciences (geography), humanities (linguistics), and from the Lundquist College of Business (accounting analytics and marketing analytics).

The other two programs in Oregon are at Oregon Institute of Technology (OIT) (in development) and Pacific University. These are not structured to have a domain emphasis, and as such are not comparable.

If applicable, explain why collaborating with institutions with existing similar programs would not take place.

Because of the long standing relationship in areas of computational genomics and biological data science between UO and OSU, there is likely to be collaboration. The two institutions are in active conversations regarding at least coordinating degrees in biological data sciences, and possibly collaborating on a joint program in this area that provides significant mutual benefits to each institution.

The other two programs at OIT and Pacific are structured in such a way that they do not present a natural area for collaboration with UO. In particular, the core aspect of domain emphases at UO, and the lack of such emphases at OIT and Pacific, makes collaboration difficult.

Describe the potential for impact on other institution's programs.

The impacts on OSU and UO are likely to be mutually beneficial. The two institutions have complementary strengths in the area of biological data science (e.g., increased model organism research occurring at UO and increased environmental life sciences interests at

OSU). Importantly, research and educational collaborations in these areas have occurred to varying degrees over the past decade, providing support for the likelihood of success of such a joint program.

Document your due diligence in consulting with other Oregon institutions.

A preliminary proposal document was submitted to the Statewide Provost's Council for review on August 26, providing an opportunity for other Oregon institutions to provide feedback. See attached documentation.

Early conversations about developing a joint program in biological data sciences between UO and OSU have occurred among Brett Tyler, director of CGRB at OSU, Bill Cresko in his role as executive director of DSI at UO, and Irem Tumer and David Conover, the vice presidents of research from OSU and UO, respectively. In addition, Director Cresko served on the external review committee for the CGRB at OSU at the request of Vice President Tumer, providing a deep understanding of the likely structure of the OSU program.

If the program's location is shared with another similar Oregon public university program, provide externally validated evidence of need.

Attach Corroborating Documentation

Please see the attached Statewide Provost's Council document

Financial Sustainability

What financial resources are needed to support this proposal? Identify the resources currently available as part of existing UO programs or reallocations within existing budgets. Are additional resources needed?

At startup, the program will require an investment of funding for 2.0 FTE administrative positions, expanding to 2.5 FTE in year two. The Data Science Initiative will provide funding for a 1.0 course creation NTTF for the first three years to support teaching faculty in the development of new courses. Other resources, including teaching hours, are available through existing UO programs. Based upon our ten year budgeting and realistic expectations of student enrollment, we project a positive revenue stream by year five of the program at the latest.

List Collaborating Units

CIS, Math, Biology, Geography, Linguistics, Accounting Analytics & Marketing Analytics

Provide a plan that shows how long-term financial viability of the program is to be achieved, addressing anticipated sources of funds, the ability to recruit and retain faculty, and plans for assuring adequate library support over the long term. (This is a header, no text needed)

Business Plan Description

The startup of the program will occur through a combination of central funds, strategic investment from the Data Science Initiative's funding, and reallocation of effort. TTF hired in affiliation with the initiative will be recruited to teach some of the lower division courses, and other courses are allocated to be taught by the domain areas. Additional staffing as listed above will be anticipated at full maturity of the program. At that time, a return on tuition via CAS mechanisms will support these functions.

We anticipate that the key partnering departments, such as CIS and math, will require additional GE resources as student growth will require expanded GE support.

Describe your plans for development and maintenance of unique resources (buildings, laboratories, technology) necessary to offer a quality program.

Nothing required at this time beyond standard university investment. At maturity, the program will require access to more small collaborative spaces, seating approximately 24-30 students than currently exist on campus. The need for this type of space will not be unique to this program and we anticipate it will align with other campus space priorities in the next ten years.

What is the targeted student/faculty ratio? (student FTE divided by faculty FTE)

At full scale, we expect approximately 45 TTF/NTTF to be involved in delivering the program, resulting in approximately a 20/1 student/faculty ratio. Data science is a highly interdisciplinary field, thus a large contingency of this faculty will include CIS and MATH TTF/NTTF.

What are the resources to be devoted to student recruitment?

The Director of Undergraduate Studies and the Advisor of Studies will work with UO Office of Admissions and Financial Aid to recruit students to the program.

Attach supporting documentation for financial sustainability.

N/A

If grant funds are required to launch the program, what does the institution propose to do with the program upon termination of the grant?

N/A

Untitled Section – Resources

List any additional faculty who will have a role in this this program as a result of the change(s), indicating those who will have leadership and/or coordinating roles. For each individual, indicate status with respect to tenure track (TT or NTT), rank, and full-time or part-time.

The development and launch of the Data Science program at the university under the auspices of the Presidential Initiative in Data Science (DSI) is a new type of endeavor for the University of Oregon. As such, the inclusion and addition of faculty to lead, support and teach in the program will occur in a different manner than would occur, for example, through the development of a new major within an existing department.

Because the major is meant to provide foundational training that will be applied through applications to domains in partnership with existing units (see program description), the development of the program primarily involves the inclusion of existing courses and faculty at UO. For example, core courses in math and CIS in applied mathematics, statistics, and computer science either already exist or were planned to be offered by faculty in those units in collaboration with the Data Science program.

In addition, over the past several years faculty lines have been approved as being affiliated with the DSI and the majority have been successful. These faculty members are planning courses that will contribute to the core offerings (e.g., data science ethics will be offered via faculty hired into philosophy in CAS) and domain offerings at the upper division (e.g., hires in biology, psychology, business, education, etc.).

As outlined in the program description, only a handful of new courses will need to be developed (e.g. DSCI 101, 102, 311, 345, 372, 411) specifically to fulfill the pedagogical needs of training in core data science principles. These courses will be developed through two primary mechanisms. First, working under the guidance of TTF members of the Internal Advisory Board (IAB), key NTT in DSI (e.g., Drs. Jake Searcy, Emily Beck, Clay Small) will work to build the courses and begin delivering content. These individuals are highly trained in data science and have been developing and teaching courses in a variety of data science areas at UO. Second, key interested TTF faculty will be recruited to participate in the development and delivery of these courses via course buyouts to home departments. Lastly, as these courses and the program grows, faculty lines will be proposed in the program through the institutional hiring plan process.

Leadership of the Data Science program will continue to occur initially (first two years) through the DSI, with Deputy Director Joe Sventek continuing to play the role of program director. Overall coordination of courses and course offerings will occur with the assistance of an undergraduate administrator and a department/office manager. Coordination of the program with the larger research and graduate educational missions of the DSI will occur under the

guidance of Executive Director Bill Cresko. As the program grows, the addition of support staff and leadership will occur organically over time. The anticipation is that at maturity - approximately 4 to 5 years - the program will have the full leadership and support staff typical of a large department in CAS.

Describe how students will be advised in the new program

The Data Science program will work closely with central advising in Tykeson Hall to provide the majority of advising to data science students. In addition, as with other large departments or programs in CAS, FTE will be built into the budget through staff and peer mentors to provide additional, data science-specific advising support for students majoring in data science. These resources will be appropriately scaled as the program grows. We anticipate that at full scale, a 1.0 FTE advising role will be provided.

What other additional staff are needed to support this program?

The program will grow organically over time, with the goal of reaching full capacity ten years after launch. At maturity, we anticipate a program that will be comparable in size to a large department in CAS, such as CIS. As such, in addition to core and associated faculty for teaching, we will require staff and personnel (see below) for office and undergraduate degree management, advising, graduate employees and technology coordinators for assisting in computational laboratories of newly developed courses.

In addition, we will require the services of an internship and capstone coordinator specifically to assist with the unique capstone aspect of our program and interactions with partners for the internships. This last component is strongly informed by our interactions with the Bioinformatics and Genomics Internship Program (BGMP) in the Knight Campus Internship Program (KCIP). During this period of growth, we will work with staff of the KCIP to provide shared services for the capstone and internship.

Full scale anticipated staffing needs:

- 1.0 FTE Department/Office Manager
- 1.0 FTE Undergraduate Administrator
- .5 FTE Director of Undergraduate Studies
- 1.0 FTE Advisor
- 33 GE Terms
- 1.0 Capstone and Internship Coordinator (3 years out)
- 1.0 Technology and Lab preparator
- Program Director – Service assignment for TTF

Are special facilities, equipment, or other resources required because of the change (e.g., unusual library resources, digital media support

Due to the UO's commitment to increasing IT infrastructure over the previous 5 years (particularly investments through Transform IT and the creation of Talapas and RACS) the overall computing infrastructure is largely in place to support our degree. We will require specialized local equipment, such as a laptop pool, an educational condo on Talapas, and support for the Jupyter collaborative coding environment.

In addition to equipment, we will require large classroom space for lectures and spaces for interactive computer education. The latter types of spaces will be our main constraint. The collaborative rooms in the Price Science Library (B042 and B044) are exemplars for this type of 'flipped' classroom but are in high demand. The opening of the KCASI will provide the possibility for sharing more of these types of rooms, the demand for these will also be high. New dry, computational spaces will need to be created to support our program at maturity.

We anticipate no unusual library needs or other extended infrastructure resources at this time.

Attach your communication(s) showing due diligence in consulting with UO Libraries and any other resource area effected by the new program.

N/A

Other Program Characteristics

Must courses be taken for a letter grade and/or passed with a minimum grade to count toward the proposed major? If so, please list the courses and the requirements of each. Although there is variation in detail, UO majors typically require that most of the courses be taken for a letter grade (not “pass/no pass”) and that the grade be C- or better.

All courses counted towards completion of the Data Science Major must be taken for a letter grade and passed with a grade of C- or better.

How much course overlap will be allowed to count toward both the major and some other credential a student might be earning (a minor, certificate, or another major)? If there are specific credentials with overlap limits, please list those and the limits.

There are no limits on course overlap counted towards another credential

Does your proposal call for new courses, or conversion of experimental courses into permanent courses? If so, please list courses in the text box below and indicate when they will be submitted to UOCC for approval:

The majority of courses that constitute the new data science degree requirements currently exist. The proposal requires the development of six new courses:

DSCI 101: Foundations of Data Science I
DSCI 102: Foundations of Data Science II
DSCI 311: Principles and Techniques of Data Science
DSCI/MATH 345: Probability and Statistics for Data Science
DSCI/CIS 372: Machine Learning for Data Science
DSCI 411: Data Science Capstone Project

Each of the above courses are concurrently submitted for approval.

Please note: DSCI 101 and DSCI/MATH 345 will be open to all students who meet prerequisite requirements. All other proposed DSCI courses will be open to DSCI majors only.

The Data Science program has solicited due diligence regarding potential overlap of DSCI 345 with all units offering statistics courses. Based on this feedback, the following restrictions will be enforced (these restrictions will also be noted on the individual course submission):

A student may not receive credit for DSCI/MATH 345 and also receive credit for any of the following courses: GEOL 418, MATH 343, MATH 461, OBA 311, and OBA 312.

Will admission to the program be limited?

No

Will students be required to apply for entry to this program?

No

Additional Requirements (Will Appear in Catalog)

Residency Requirements (Will Appear in Catalog)

At least 34 credits of coursework applied to the major must be taken at the University of Oregon. These credits must include enrollment in DSCI 311, DSCI 345, and DSCI 372.