Utah System of Higher Education
New Administrative Unit Proposal
Cover/Signature Page - Abbreviated Template

Institution Submitting Request:	**School of Computing**
Proposed Effective Date[1]:	**10/01/2019**
Institutional Board of Trustees' Approval Date:	
Proposed Unit Title:	**Utah Center for Data Science**
Sponsoring School, College, or Division:	**College of Engineering**
Sponsoring Academic Department(s) or Unit(s):	**School of Computing, Department of Mathematics**

Proposed Unit Type:

☐	New Administrative Unit
☐	New Center
☐	New Institute
☐	New Bureau
☒	Conditional Three-Year Approval for New Center, Institute, or Bureau

Chief Academic Officer (or Designee) Signature:
I, the Chief Academic Officer or Designee, certify that all required institutional approvals have been obtained prior to submitting this request to the Office of the Commissioner.

Jeff M. Phillips Date: June 13, 2019

☒ I understand that checking this box constitutes my legal signature.

[1] "Proposed Effective Date" refers to date after Regent approval when new unit is operational or change to unit is published.

New Unit Description - Abbreviated Template

Section I: The Request

School of Computing requests approval to establish Utah Center for Data Science effective 10/01/2019. This action was approved by the institutional Board of Trustees on .

Section II: Program Proposal

Administrative Unit Description/Rationale

Present a brief description of the unit. Describe the institutional procedures used to arrive at the action being proposed. Briefly Indicate why a new administrative unit or change to the unit is justified. Are similar units offered elsewhere in the USHE or the State? State how the institution and the USHE benefit from the proposed unit or unit change.

Overview:

The data science faculty at the University of Utah request the formation of the Utah Center for Data Science (UCDS). Data science is an interdisciplinary field, which spans topics in computing, data management, and statistical analysis. The University of Utah has experts in these areas spanning Computer Science, Mathematics, as well as many other departments in Engineering, Sciences, Business, and Humanities. Their collective research has benefitted from cross-department and cross-college collaboration, typified by work involving a domain expert and a computing/management/analysis expert. Through these academic departments, a number of data science related educational programs have recently been born. These programs have largely grown independently, although there is a growing sense of coordination among them. Other recent events members of the proposed center have led include a Data Science Day, a newly formed student-driven Data Science Club, and joint research seminars. Momentum for data science is evident in the state and throughout the country; this request addresses the need to centralize and organize these activities in Utah.

On Data Science:

Data Science is an evolving, but quickly developing field finding many applications across disciplines. In October 2015, the American Statistical Association released a statement on the Role of Statistics in Data Science, and attempted to summarize the field:

"...its interdisciplinary nature means that a substantial collaborative effort is needed for it to realize its full potential for productivity and innovation. While

there is not yet a consensus on what precisely constitutes data science, three professional communities, all within computer science and/or statistics, are emerging as foundational to data science: (i) Database Management enables transformation, conglomeration, and organization of data resources, (ii) Statistics and Machine Learning convert data into knowledge, and (iii) Distributed and Parallel Systems provide the computational infrastructure to carry out data analysis."

One can also categorize research around data science into two camps:
(1) research *in* data science: developing new techniques and methodology in how to more efficiently and effectively perform data-first science.
(2) research *using* data science: working at the forefront of many scientific and engineering tasks, by examining and drawing conclusions through the analysis of topical data sets.
Data science is a new mechanism for advancing science, engineering, and society. Researchers in data science build and extend this machine, and researchers using data science often work with those in the first camp to tune this mechanism to their specific goals.
The proposed center will be built around researchers in the narrower, but farther reaching camp of research *in* data science. Its goal will be to train and organize those in the former camp, but also provide a resource and aid for the large and growing researchers and students in the latter camp (of *using* data science).

At Utah, data science (whether *in* or *using*) has grown to be a part of many faculty's core research, and through a variety of academic programs. However, these efforts are currently only loosely organized and coordinated. To address this, the proposed Utah Center for Data Science will have the following mission:

* Provide a central structure to lead, manage, and organize data science researchers and resources at the University of Utah.
* To establish and advance through research and applications the fundamental principles and practice of data science.

There currently does not exist such an entity. Presently, we plan to achieve this along 4 vectors:

* Coordinating research among data science specialists, and providing a central location for other researchers to find the appropriate collaborator.
* Organizing outreach events to engage the university and broader Utah community.
* Overseeing student engagement in data science, centered on the new Data Science Club.
* Enhancing and growing a research program in foundations in data science, starting with a joint seminar and joint research efforts.

The UCDS plans to use the web domain datascience.utah.edu to provide a front-facing location for achieving this mission. We already host several proposed projects that would fall under the center at this site.

Coordinating Cross-disciplinary Research Activities:
Although data science is becoming a central part of many disciplines around campus, there is no central place for researchers to find the ``right" collaborator for a particular data science project they are pursuing. As described above by the ASA statement, data science has several sub-communities, and within each of these sub-communities members of the proposed center have different expertise. By collating information about members and providing a useful interface on the site, it will greatly facilitate this process. Such collaborations have already led to much productive research and numerous grants of external funding (including from the data science specific NSF BIGDATA call, as well as many domain specific calls within NSF, DARPA, and DOE enhanced by data science components). We expect the center heading and planned organization will help facilitate even more efforts and output.

Community Outreach:
The center will help coordinate events that bring together many members of the university and broader Utah community. An initial event, the January 2017 Data Science Day, with 163 registered participants (and more unregistered), was organized by the proposed members and through the website http://datascience.utah.edu/dataday/. It included a job fair, an industry panel, a poster and demo session, invited talks from four faculty around campus, and a

keynote from a prominent data scientist at Amazon. The center would plan to make this an annual event (moving to Falls in 2018), and expand its scope and engagement. The center also hope to use its purview to engage with other local data science related community groups such as the Utah Big Data User's Group and the Utah chapter of Lambda Lounge.

Data Science Club:
Utah Center for Data Science members have recently help guide the creation of a student-led Data Science Club, launched in Fall 2017. The initial meeting with a panel of local data scientists had about 100 students from across campus in attendance. Bi-weekly meetings have consistently had at least 25 students; these sometimes have food sponsored by local companies eager to interact with and recruit these students. The food greatly increases involvement. The center will plan to sponsor this club underneath its banner to help provide consistent food, coordinate activities, speakers, and job opportunities for the club's members and participants. Additional events in this space may include hack-a-thons and field trips to local companies.

Coordinating Foundational Data Science Efforts:
Beyond the inter-disciplinary efforts described above, the center hopes to tighten the collaboration among core members in the foundation of data science. This will hopefully increase opportunities for many funding calls specific to the foundations of data science, such as the recent NSF TRIPODS program, the NSF BIGDATA program, and those within the NSF CISE-III division.
Also, while there are numerous seminars around campus which touch on data science (including several in the Computer Science and Math Departments) these are not always well-coordinated with all researchers who would like to attend. The center will plan to run a central, cross-department and cross-college, Data Science Seminar, that will cover diverse topics in data science from foundations to applications, and be open to all across campus. To make this seminar a larger event, UCDS funding will help support food, and occasional invited speakers from outside the university.

Data Science is truly a national (and international) growth area in industry and funding. It is one of the fastest growing area for job opportunities (and with high salaries) and for external funding opportunities. For instance, the National Big Data Research and Development Initiative featured more than $200 million in new funding commitments from six government funding agencies. Many other universities are creating large data science centers or institutes (for instance, University of Michigan, University of Washington, and NYU to just name a few). The establishment of this center will help ensure that the University of Utah's university-wide effort in data science is the strongest in the Mountain West. This will benefit the University of Utah and the entire USHE in attracting strong students and businesses to its programs and to the state.

This proposal is a joint effort by the data science faculty of the University of Utah, all of whom have read, contributed to, and approved this proposal. Creation of a Center for Data Science has been in discussion for several years, and this proposal is the culmination of several discussions about how as a group Utah's data science faculty can best serve the university and broader Utah community. In this time the core members have developed a report and sense of place through overlapping seminars, joint research and grants, co-advising students, co-teaching classes, organization of the Data Science Day, and discussions about curricular issues.

This is especially true among the "Data Group" consisting of 5 proposed center members. This has grown from a more focused group on algorithm and databases to include many aspects of data science. Most of the other proposed center members and affiliate members have spoken at the seminar. As the number of faculty working in the foundational areas of data science has grown and intersected, the structure has grown beyond that of a joint research group, and is ready to take on the expanded role of a center.

Additionally the University of Utah supported a Transformative Excellence Program (TEP) in Statistics between the Math Department, the School of Computing, and the Department of BioEngineering. The faculty most involved in this effort, and faculty hired as a result of this TEP are proposed members. This center would be a primary result of the collaborations and hiring spurred

by the TEP.

A few somewhat similar centers exist or existed within USHE, specifically at the University of Utah. But these centers are mainly focused on the research of a particular group:

* Center for Extreme Data Management and Visualization: http://cedmav.org
  led by director Valerio Pascucci, focuses primarily on the scalability issues within scientific data
* Global Knowledge Management Center: http://gkmc.utah.edu
  led by director Olivia Sheng, focuses primarily on the knowledge representation for information systems and business applications

Additionally other centers and institutes focus on particular type of data (e.g., sociological, imaging), and largely consist of members using data science -- or have members overlapping the UCDS. In other cases, we plan on having representatives from these efforts around campus as Affiliated Members in UCDS to help ensure these missions are coordinated.

Consistency with Institutional Mission/Institutional Impact

Explain how the unit is consistent with the institution's Regents-approved mission, roles, and goals. Describe how the existing administrative structures support the proposed unit and identify new organizational structures that may be needed. What changes in faculty and staff will be required?

The Utah Center for Data Science will provide important organizational support and credentials to Utah's efforts in research, education, and outreach.
Members of the center already maintain the datascience.utah.edu web domain, and will continue to expand its contents.
Most of the activities are already taking place under the faculties normal duties, and no additional faculty or staff are required.

The center will have a Director who heads the center, and will report to the Dean of the College of Engineering. There will also be Associate Directors of Research (in charge of coordinating the seminar), of Outreach (coordinating joint efforts on the Data Science Day and other endeavors), and of Student-Involvement (mentoring the Data Science Club and other activities). The director will be appointed by the Dean of the College of Engineering. Each associate director will be elected by the Core Members to one year terms, and

will typically serve for about 3 years. The specifics on these roles and the election procedure are described in the attached policy document "Governance Roles for Utah Center for Data Science."

Core Members are researchers *in* data science, and will be the ones tasked with running the center. Initial Core Members will include:
Associate Professor Jeff M. Phillips (Computer Science),
Professor Suresh Venkatasubramanian (Computer Science),
Associate Professor Feifei Li (Computer Science),
Assistant Professor Vivek Srikumar (Computer Science),
Assistant Professor Bei Wang Phillips (Computer Science),
Professor Lajos Horvath (Mathematics),
Assistant Professor Braxton Osting (Mathematics),
Assistant Professor Alexander Lex (Computer Science),
Assistant Professor Aditya Bhaskara (Computer Science),
Assistant Professor Shandian Zhe (Computer Science)

The center will also include Affiliated Members who are close collaborators of the core members and are leaders in research *using* data science. Their main role will be keeping the UCDS connected with the uses of data science around the university. Initial Affiliated members will include:
Professor Ken Smith (UPDB),
Associate Professor Kyle Dawson (Physics and Astronomy),
Professor Olivia Sheng (Information Systems),
Associate Professor Zac Imel (Educational Psycology),
Professor Sarang Joshi (Biomedical Engineering),
Associate Professor Tolga Tasdizen (Electrical Engineering),
Professor John Horel (Atmospheric Science),
Assistant Professor Xiaoyue Cathy Liu (Civil and Environmental Engineering),
Assistant Professor Amanda Smith (Mechanical Engineering),
Assistant Professor Taylor Sparks (Materials Science and Engineering),
Associate Professor James Sutherland (Chemical Engineering),
Associate Professor Brian Chapman (Radiology, Biomedical Informatics)

The Initial Core Members of the proposed center have unanimously agreed to a

policy by which new Core Members and Affiliated Members can join. Twice a year they will process internal nominations for members and external requests for affiliated membership. New Core Members will require at least 2/3 voting in agreement from current Core Members, and new Affiliated Members will require above 1/2 of Core Members voting in agreement. The attached policy document "Membership Plans for Utah Center for Data Science."

The administrative staff at the School of Computing will initially provide all needed administrative support associated with budgets, room and food bookings, and travel arrangements. No addition faculty or staff are needed for explicit connections to the center.

The official center designation will increase visibility of the strengths of Utah's data science efforts within the university and outside of it. We also except this more dedicated and organized focus to lead a continued growth in data science with respect to funding, educational output, and community engagement.

For Conditional Approval of New Centers, Bureaus, and Institute's ONLY

Explain how proposed change is fully supported by external funding.

All center activities have been operating with almost no budget (some industrial support). The ongoing research activities are sponsored by the Core Member's active research grants, and will continue to be so. We have applied for governmental support (NSF, pending) and believe that having a formal center will aid in finding additional government and industrial sponsorship. We expect to find additional funding before transitioning to a permanent center, but also believe the center is capable of being run successfully without substantial new funding.

Explain how the proposed change requires only a modest effort in terms of staff and space needs with normally no permanent staff or permanent facility assignment. Explain how the entity is already affiliated with an existing academic program or department.

The center is a virtual center with no explicit space requirements, although the SoC director has agreed that the lab space (MEB 3445) already used by many of the Core Members research groups can be a hub for UCDS activities. The center activities have been running using existing resources already. SoC and CoE (where most Core Members are affiliated) have agreed to continue to support the UCDS with any minor administrative support required for activities.

Finances

What costs or savings are anticipated with the actions proposed? What new facilities or modifications to existing facilities or equipment are needed? Describe any budgetary impact on other programs or units within the institution. If new funds are required, describe expected sources of funds.

The Utah Center for Data Science will use promised support from the Department of Mathematics, School of Computing, and College of Engineering. In particular, each have promised seed funding of $5000, for a total of $15000. The intent of this funding is to provide initial support for the operational costs of

a larger and more prominent seminar, the data science day, and other organizational events.

The Utah Center for Data Science will solicit activity-specific support from local industry. For instance, many local companies sponsor pizza during presentations at the Data Science Club, and we will seek sponsorship to help support the future Data Science Days. We believe industrial sponsorship will eventually be able to entirely support all of the aforementioned engagement and outreach activities.

Members will also write for additional support within future external funding proposals that explicitly contributes to the center's budget. These funds will contribute to support staff--technical and administrative--which will further help amplify the center's impact.

Several of the members of the proposed center already locate their research groups in a large, recently remodeled lab in the Merrill Engineering Build (MEB 3345), controlled by the School of Computing. The SoC Director, Ross Whitaker, has promised that this space can be designated as a hub for center activities. Some of the faculty in the Department of Mathematics will locate in the new "Data Science Center" space in the new Crocker Science Center. No additional space will be required, as in most ways the UCDS will be virtual.

Letters of Support, including those pledging the specific above financial and space support, are attached from:

* Rich Brown, Dean of College of Engineering
* Henry White, Dean of College of Science
* Cynthia Berg, Dean of College of Social and Behavioral Science
* Taylor Randall, Dean of David Eccles School of Business
* Ross Whitaker, Director of School of Computing
* Davar Khoshnevisan, Chair of Department of Mathematics
* Wendy Chapman, Chair of Department of Biomedical Informatics