UNIVERSITY OF CALIFORNIA, ACADEMIC SENATE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

Susan Cochran
Telephone: (510) 987-0887
Email: susan.cochran@ucop.edu

Chair of the Assembly of the Academic Senate
Faculty Representative to the Regents
University of California
1111 Franklin Street, 12th Floor
Oakland, California 94607-5200

February 22, 2023

KATHERINE S. NEWMAN
PROVOST AND EXECUTIVE VICE PRESIDENT
UNIVERSITY OF CALIFORNIA

Re: Approval of Master of Data Science in Health (MDSH) at UC Los Angeles

Dear Katherine:

In accordance with the *Universitywide Review Processes For Academic Programs, Units, and Research Units* (the "Compendium"), and on the recommendation of CCGA, the Academic Council has approved UC Los Angeles's proposal to establish a Master of Data Science and Health (MDSH) self-supporting graduate and professional degree program (SSGPDP).

Because this is a new degree title, and the Assembly of the Academic Senate is not meeting within 30 days of CCGA's approval, Council must approve the program per Senate Bylaw 125.B.7.

I am enclosing CCGA's report on its review of the new program, and respectfully request that your office complete the process of obtaining the President's approval.

Sincerely,

Susan Cochran, Chair
Academic Council

Cc: Academic Council
IRAP Analyst Procello
UCLA Senate Director de Stefano
Executive Director Lin

UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

COORDINATING COMMITTEE ON GRADUATE AFFAIRS (CCGA)
Erith Jaffe-Berg, Chair
erithj@ucr.edu

ACADEMIC SENATE
University of California
1111 Franklin Street, 12th Floor
Oakland, California 94607-5200

February 6, 2023

ACADEMIC SENATE CHAIR SUSAN COCHRAN

Dear Chair Cochran:

Following the Academic Council discussion on January 25, 2023, I am submitting this revised letter about the Master of Data Science in Health program at the Los Angeles campus (MDSH).

The Coordinating Committee on Graduate Affairs (CCGA) fully reviewed this proposal, with four external reviewers, a lead reviewer from the CCGA, discussions and a vote at the CCGA. The CCGA voted unanimously (with one abstention) at its December 5 meeting to approve the MDSH proposal. As with all proposals, the CCGA lead reviewer's report was attached to my communication to the Senate. The CCGA lead reviewer's report highlighted some concerns raised by the external reviewer, which had been raised in communications with the program proposers. At the Academic Council meeting on January 25, concerns in three areas were raised. As a result, the proposal, originally on the consent calendar, was rescinded from the calendar.
Subsequently, I brought the issue before CCGA members, and we discussed the concerns. Based on those discussions and a review of the information available, I affirm the decision originally taken by the CCGA at its vote on December 5. I am presenting the proposal once again here, with an expanded letter, addressing concerns raised at the Academic Council.

The CCGA determined that the proposed Master of Data Science in Health (MDSH) will meet the demands of the health industry and increase the number of well- trained health data analytics leaders.

The MDSH program will cater to prospective students seeking graduate-level professional education in health analytics who are (i) working professionals, (ii) residing geographically farther than UCLA's current, state-funded programs, and (iii) recent domestic or international college graduates seeking an advanced degree for entry to the data science in health industry. MDSH courses will be offered and delivered in a hybrid mode (weekend and online classes) so that professionals can complete the degree conveniently with minimal disruption to normal business hours

The program will be designed to help address the acute shortage of trained health data scientists in California and the nation. With its emphasis on data science, the curriculum will offer coursework that brings academic rigor to healthcare and public health professionals in data engineering, data visualization, data mining and exploring, machine learning, and research design.

Four reviewers (plus the Lead Reviewer) evaluated the proposal. All reviewers stated a general agreement with the rationale and broad learning objectives of the program. Consensus is that there will be a high demand for the program and excellent career opportunities for its graduates. The reviewers did identify some weaknesses and concerns, which are addressed in the lead reviewer's report (attached).

A concern was raised by the UCPB related to the fact that a Master of Data Science in Biomedicine at the UCLA Geffen School of Medicine was approved last year. UCPB wanted assurance that the proposers of the current program were aware of this other program and were acting in coordination. This issue was also raised in the CCGA lead reviewer's report and in an external reviewer's report. CCGA has had more communications with the campus regarding this concern. The MDSH proposal was developed and approved by the UCLA Graduate Council before the Department of Computational Medicine's Master of Science in Data Science in Biomedicine was approved. However, the MDSH was A new degree designation and needed approval by the Legislative Assembly. Due to the timing of the Legislative Assembly it was not routed to the CCGA until the end of the spring quarter, which meant the other program was approved first. In comparison, the programs, though apparently similar, have important differences. The MDSH is a two-year, 48 unit hybrid program. The MS in Data Science in Biomedicine is a 36 unit online program. In terms of their curricula, the Chairs of the departments of Computational Medicine and Biostatistics submitted a joint memo outlining the differences between the two programs and the potential contribution each program has to the field of data science. The memo highlights that though there are shared elements related to training in computational statistics and machine learning, and some shared elective courses, the programs are directed at different applicant populations. The MDSH is focused on those already in careers (or those seeking a career in) statistical modeling and studies in clinical data science research or in pharmaceutical companies, insurance companies, public health research and other data technology companies. The program approved last hear is aimed at professionals in engineering or data scientists in BIG DATA technologies deployed by pharmaceutical, bioinformatics and biotechnology companies.

Another concern raised by UCEP pertained to the potential impact on undergraduates if faculty are diverted from state-supported to self-supported programs. The CCGA lead reviewer affirmed that the Data Science committee and the faculty at UCLA voted unanimously in favor (19/23 votes 19 in favor) of approving the program, fully aware of the teaching requirements of the program. The UCLA Graduate Council furthermore approved the program with awareness of the campus impact of the teaching needs. Curricular issues about course preparedness were alluded to by the UCEP at the Academic Council, and those were also raised and brought to the attention of the proposers in the lead reviewer's detailed report.

The UCAADE Chair raised concerns about diversity. Our lead reviewer also highlighted that the plans to return to aid of 10% is relatively low. The program does note: "to ensure that the program recruits a diverse pool of students, especially the underrepresented and/or marginalized communities, for the first cohorts (before the program achieves solvency), the Department of Biostatistics will commit 50% tuition and fee coverage for two entering students in the first cohort (estimated fund $58,320) and one entering student in the second cohort (estimated fund $29,889)." (p.19). Diversity, equity and inclusion are areas for the campus to continue to monitor in its third year review of the program, and the CCGA report emphasizes this.

As you know, CCGA's approval is the last stop of the Academic Senate side of the Systemwide review and approval process except when the new degree title must be approved by the Academic Council. I submit this for your review and have enclosed the Lead Reviewer's report. Please do not hesitate to contact me if you have further questions regarding the proposal.

Sincerely,

Erith Jaffe-Berg
CCGA Chair

cc: James Steintrager, Academic Council Vice Chair
CCGA Members
Monica Lin, Academic Senate Executive Director
Michael LaBriola, Academic Senate Assistant Director
Chris Procello, Academic Planning and Research Analyst

CCGA Review of Program Proposal

„Master of Data Science in Health (MDSH) Self-Supporting Graduate Professional Degree Program (SSGPDP)“

Frithjof Kruggel (UC Irvine, Lead Reviewer)

Program Leads and Location

1. Sudipto Banerjee (UC Los Angeles, sudipto@ucla.edu)
   Professor and Chair of the Department of Biostatistics

2. Hua Zho (UC Los Angeles, huazhou@ucla.edu)
   Professor, Committee Chair, Biostatistics Department Data Science Committee

Department of Biostatistics, Fielding School of Public Health (FSPH)
University of California, Los Angeles (UCLA)

Proposal Outline

The proposal consists of 995 pages, including 10 pages of cover letters, 27 pages of proposal description, and 23 appendices, among them a market analysis and 809 pages of Faculty CVs. Structure and content are complete and adhere to the format as outlined in the CCGA Handbook Appendix B and J, and the Policies for Self-Supporting Graduate Professional Degree Programs (Sep 2020).

The proposal describes a SSGPDP leading to a MS Degree of Data Science in Health, following a Master's II plan with a capstone project. The program consists of eight required and four elective courses (chosen of seven offerings), with a total of 48 units. Students will be admitted on an annual basis, starting in Fall. Students enroll in one to three courses per quarter, with a normative time-to- degree of six academic quarters and a maximum of five years. Topics of the required courses include (1) one course in Public Health, (2) three courses in Data Science, (3) three courses in Data Analytics, and (4) one capstone course. Elective classes focus on Biostatistics and are offered by the proposing Department. The course PUBHLT C201 „Fundamentals of Public Health“ is given fully online. All other courses will be taught in hybrid mode, consisting of three weekends of in-person sessions (24 hours) and eight online evening sessions on weekdays (16 hours). „Hybrid mode“ is defined here as a „non-trivial combination of (i) in-person, (ii) online, and/or (iii) remote primary offerings“. The targeted audience includes prospective students seeking an advanced degree for entry to a career in data science in the health industry, namely: (i) working professionals; (ii) students residing outside of UCLA's reach; and (iii) international college graduates.

Chronology of the Review Process

Aug 2020	Data Science Committee formed
May 2021	Faculty unanimously voted in favor (19/23 votes, 19 in favor)
Jan 2022	Review and approval by UCLA's Graduate Council
June 2022	Proposal submitted for systemwide review
Oct 12, 2022	Proposal assigned for CCGA review
Nov 4, 2022	UCPB review submitted
Dec 5, 2022	Discussion at CCGA meeting; proposal approved with modifications

Reviewers

1. Patrick Heagerty (U Washington, heagerty@uw.edu) (proposed by program leads)
   Professor and former Chair of the Department of Biostatistics. He is the Director of the Biostatistics and Research Design Core for the NIH Health Care Systems Research Collaboratory, for the NIH Mental Health Research Network, and a member of the Executive Committee for the FDA Sentinel Innovation Center.

2. Nick Jewell Biostatistics (UC Berkeley, jewell@berkeley.edu) (proposed by program leads)
   Professor of Biostatistics and Statistics at the School of Public Health. He is a member of the National Academy of Medicine, and was Vice Provost of the UC Berkeley campus, and later at the Office of the President. His appointment is joint between Statistics and Biostatistics in the School of Public Health.

3. Eric B. Sudderth (UC Irvine, sudderth@uci.edu)
   Professor of Computer Science and Statistics, and Chancellor's Fellow. He directs the UC Irvine Center for Machine Learning and Intelligent Systems, as well as the HPI Research Center in Machine Learning and Data Science at UC Irvine.

4. Kai Zheng Informatics (UC Irvine, kai.zheng@uci.edu)
   Professor of Informatics and Emergency Medicine. He also serves as Chief Research Information Officer of the Office of Data and Information Technology, Director of Biomedical Informatics of the Institute for Clinical and Translational Science, and Director of the Public Health Informatics & Technology Workforce Development Program.

Because this program is self-supporting, a con-commitant review by the Committee on Planning and Budget was solicited.

Strengths of the Proposed Program

All reviewers stated a general agreement with the rationale and broad learning objectives of the program. Consensus is that there will be a high demand for the program and excellent career opportunities for its graduates. Reviewers agree that size and expertise of faculty are adequate to administer the program. The program structure is comparable to those of other programs (e.g., Harvard University, U NC, Chapel Hill, U Washington, U Michigan, UC San Francisco). Course content was deemed to be excellent (Reviewer 2). Admission criteria, DEI considerations, program timeline and evaluation plans are adequate. CPB considered the demand for the program as high and its cost projections as realistic. Profits should be sufficient to compensate the use of campus resources.

Weaknesses and Concerns

Overall, weakeness and concerns discussed below are considered minor. However, their discussion will be given sufficient room so that the program can clarify, respond and amend as needed.

1. *Similar programs:* Recently, a program „Master of Data Science in Biomedicine“ at the UCLA Geffen School of Medicine was approved. This proposal is not referenced or discussed here. Concerns were raised that two programs with similar content and audience but different program requirements and costs will compete for the same student pool. It is suggested that this programs analyzes similiarities and differences of both programs, with the option of clarifying and sharpening its profile and/or discuss cross-listing courses. This may help to advise prospective students, and help to avoid possible confusion.

2. *Prerequisites:* Admission requirements (p. 12) declare that „undergraduate coursework in differential and integral calculus and elementary linear algebra is preferred but not required“, yet Appendix C (p.

33) states such coursework as a requirement. How will students without that background handle the necessary courses? It is strongly suggested to declare this coursework as required. Given the potential diverse audience, prerequities in terms of computer literacy should be clarified.

3. *Diversity:* It is expected that diversity is similar to other programs, with a high proportion of female students, but a low respresentation of URM students. Only 10% of annual revenue is allocated for need-based scholarships which is considered as the minimal acceptable rate. Which efforts will the program undergo to attract and support URM students?

4. *Hybrid presentation mode:* Reviewer 4 argued that in-person exposure in teaching is critical. It is suggested that the program critically reviews the benefits and success of this presentation mode in its first three-year review, with the option of reverting to in-person mode.

5. *Coursework:* BIOSTAT 100 appears to be an upper-division undergraduate course. Is this by design or oversight? The program should consider adding (or cross-referencing) coursework in computer science/programming so that students are prepared and can focus on the content provided here.

6. *Breadth in Vision:* It is understood that the program will solely be governed by Faculty of the Biostatistics Department. At the same time, the proposal clearly argues that "data science draws upon multiple disciplines" (p. 17). It would be advantegeous to include greater representation of fields outside Biostatistics, in order to provide students with a broader view of application areas and career opportunities. In this context, it is suggested to provide students a more in-depth knowledge of the nature of heterogeneous health data sources.

7. *Teaching Assistants:* "will be hired from the MS, MPH and PhD student pool" (p. 23). Do these students have sufficient training in Data Science to staff all courses? Will there be a training program for prospective TAs?

8. *Accreditation:* The Public Health course was included as a requirement for the Council on Education for Public Health (CEPH) accreditation of this MS program. What is the value of the accreditation to the program and its alumni?

9. *Computing resources:* As stated on p. 22 of the proposal, it is not expected that „any additional resources will be needed" and that „cloud and cluster computing cost" will be covered within the instruction budget. The proposal also states (p. 19) that the department „plans to invest in more computing nodes in the UCLA Hoffman2 cluster." Who is paying for this investment?

10. *Campus resources:* The use of facilities is not accounted for as a direct cost, but a modest campus tax of 5% on revenue that might not compensate the use of campus resources. Realistic costs for weekend use of classrooms, their infrastructure and computing resources should be estimated and included in the budget.

Recommendation

It is recommended to approve the proposal with modifications agreed upon by CCGA, the program proponents, and the campus.