SPECIAL REPORT

of the

ACADEMIC PRIORITIES, GRADUATE AND
PROGRAM AND BUDGET COUNCILS

concerning a

CERTIFICATE IN GENOMIC DATA SCIENCE

Presented at the
758th Regular Meeting of the Faculty Senate
April 28, 2016

COUNCIL MEMBERSHIP

ACADEMIC PRIORITIES COUNCIL

Faune Albert, Richard Bogartz (Chair), Nicholas Bromell, Elizabeth Chilton, Suzanne Daly, Kathleen Debevec, Jean DeMartinis, Piper Gaubatz, Bryan Harvey, Masoud Hashemi, Deborah Henson, A Yemisi Jimoh, Sangeeta Kamat, Stephen Magner, Ernest May, Katherine Newman, MJ Peterson, Monroe Rabin, James Rinderle, Peter Stern, Jack Wileden, Donna Zucker

GRADUATE COUNCIL


PROGRAM AND BUDGET COUNCIL

ACADEMIC PRIORITIES COUNCIL

The Department of Biostatistics and Epidemiology in the School of Public Health & Health Sciences at the University of Massachusetts Amherst proposed to offer a Certificate in Genomic Data Science. The courses are currently offered through classroom learning (i.e., in person, not online). The proposed Certificate in Genomic Data Science is a four-course (12 credit) certificate that includes graduate-level courses from the Departments of Biostatistics & Epidemiology, Computer Science, and Environmental Health Sciences. Students can choose from an approved list of courses - the topics of the courses include statistical genetics, stochastic models in population genomics, data science and visualization, statistical learning, advanced epidemiologic methods, and other methodological courses in biostatistics. The list of approved courses is extensive so that students can choose courses that align with their interests and strengths. All approved courses include skills that are needed by potential employers.

The Academic Priorities Council voted to recommend approval of the proposal at its meeting on March 24, 2016.

GRADUATE COUNCIL

Through an electronic vote conducted on March 2, 2016, the Academic Standards and Curriculum Committee (ASCC) of the Graduate Council voted to approve the Certificate in Genomic Data Science.

On Wednesday, March 9, 2016, the Graduate Council unanimously approved the Certificate in Genomic Data Science, Proposal #2319 in the Course and Curriculum Management System.

PROGRAM AND BUDGET COUNCIL

The Program Subcommittee of the Program and Budget Council met on February 10, 2016, reviewed the Certificate in Genomic Data Science proposal and recommended it for approval.

At its meeting on February 17, 2016, the Program and Budget Council unanimously approved the Certificate in Genomic Data Science, Proposal #2319 in the Course and Curriculum Management System.

MOVED: That the Faculty Senate approve the Certificate in Genomic Data Science, as presented in 27-16 Sen. Doc. No. 16-055.
Briefly describe the Certificate.

The Department of Biostatistics and Epidemiology in the School of Public Health & Health Sciences at the University of Massachusetts Amherst will offer the proposed Certificate in Genomic Data Science. The courses are currently offered through classroom learning (i.e., in person, not online).

The proposed Certificate in Genomic Data Science is a four-course (12-credit) certificate that includes graduate-level courses from the Departments of Biostatistics and Epidemiology, Computer Science, and Environmental Health Sciences. Students can choose from an approved list of courses - the topics of the courses include statistical genetics, stochastic models in population genomics, data science and visualization, statistical learning, advanced epidemiologic methods, and other methodological courses in biostatistics. The list of approved courses is extensive so that students can choose courses that align with their interests and strengths. All approved courses include skills that are needed by potential employers.

Due to the increasing prevalence of big genomic datasets, there is an urgent need in industry and academia for workers who are trained with applications of such data in biomedical research and personalized medicine. A recent McKinsey report claims that “by 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions.” The purpose of the Certificate in Genomic Data Science is to prepare students to meet the need for qualified individuals trained in handling genomic data within healthcare and drug discovery. Acquiring the certificate will help students become more competitive in the job market; the certificate also provides concrete demonstration by the student of having obtained a set of marketable skills in genomic data science with applications to biotechnology and the pharmaceutical industry.

The required courses for the certificate are at the level of 500 and above. The certificate is open to qualified undergraduate and graduate students.

Matriculated students and non-matriculated Continuing and Professional Education students can enroll in the required courses for the Certificate in Genomic Data Science, as long as they have the prerequisites for the courses. Courses are offered in both fall and spring. The certificate can be completed within one year.

Provide a brief overview of the process for developing this Certificate.

The faculty in Biostatistics and Epidemiology determined that there is an urgent need for students trained in genomic data science at the level of certification, based on feedback from potential employers. After meeting with faculty department representatives from Biostatistics, Isenberg School of Management, Computer Science and Mathematics and Statistics, the decision was made to have three separate certificates offered by 1) Biostatistics, 2) Isenberg School of Management and 3) Mathematics and Statistics and Computer Science.

The three separate certificates are needed based on the different goals of students in each of these fields. The faculty from Biostatistics and Epidemiology met to determine the appropriate courses needed for training in genomic data science, based on industry standards and needs of potential employers. A final list of courses was identified that will be offered consistently and meet the requirements of a Certificate in Genomic Data Science.

Describe the Certificate’s purpose and the particular knowledge and skills that will be acquired by participating students.

The Genomic Data Science certificate prepares students for data-driven decision making, and deriving insights from genomic data sets using the most up-to-date statistical methods, computational methods and analytical tools available, with an emphasis on applications in biomedical research. The program blends topics in statistical methods, statistical computing, machine learning and algorithm development relating to genomic applications in biomedical research, to train students to become effective data scientists. The course list includes courses from the Department of Biostatistics and Epidemiology. Additional skills that students will develop include the ability to work with large genomic databases, to manage and evaluate data sets, and create meaningful output that can be used in effective decision-making.
If this proposal requires no additional resources, say so and briefly explain why. If this proposal requires additional resources, explain how they will be paid for. For proposals involving instruction, indicate how many new enrollments are expected and whether the courses have room to accommodate them.

This proposal requires no additional resources as all of the listed courses are offered to support existing academic programs.

Please describe the curriculum for this certificate, listing all required courses and possible electives, any prerequisites or GPA requirements, the recommended order or coursework and any other pertinent information. You may attach additional materials related to the curriculum at the end of this section.

Requirements for the certificate: Students must complete the required courses with a grade of "B+" or better for each course. All courses must be taken through UMass Amherst.

The following curriculum is subject to change when appropriate by consent of the Department of Biostatistics and Epidemiology. For prerequisite information, students can see the list of prerequisites for each individual course.

A total of four 3-credit courses listed below are required.

One or both of the following:
BIOSTATS 690MS Stochastic Models for Population Genomics
BIOSTATS 690T Applied Statistical Genetics

At least two of the following:
BIOSTATS 690DS Data Science: Visualization, Statistics and Computing
BIOSTATS 690JQ Applied Statistical Learning
BIOSTATS 740 Mixed models and analysis of longitudinal data
BIOSTATS 697G Bayesian Computation in Biostatistics
EPI 737 Intermediate Epidemiological Methods
EHS 600 Molecular Epidemiology
CMPSCI 589 Machine Learning

Explain how these courses represent a coherent course of study.

The study of Genomic Data Science includes developing in-depth knowledge of both statistical theory and computation techniques, as well as a strong foundation of the biological principles underlying genomic data. The choices of courses include training in the science of genomics, computational skills, algorithm development, machine learning and data visualization, all with a strong emphasis of applications of genomic data in biomedical research. By combining these diverse skills into one certificate, students will naturally build knowledge in data science, while also building a strong foundation in the growing field of genomic applications in biomedical research.

Describe how there is a clear educational objective that can be achieved in an efficient and well-defined manner.

The educational objectives of the Certificate in Genomic Data Science are for students to build a fundamental understanding of the computational challenges in genomic sciences, and obtain the statistical and computational expertise necessary to help address some of these challenges. The required four courses for the certificate can be completed within two semesters.

Explain how the course sequence offers a clear objective at the appropriate educational level.

All of the required courses are at the level of 500 and above, and are thus appropriate for the certificate at the graduate level, or for exceptional undergraduate students with the appropriate background.
Describe the perceived need for this Certificate.

Due to the vast amounts of genomic data that are being collected, making use of information has become key in many areas relating to genomics, especially drug discovery and personalized medicine. There is a high demand for data scientists with a focus on genomic data and this certificate would help fill that demand.

If the courses that comprise the Certificate have been or currently are being offered, describe their schedule of availability. If the Certificate is comprised of new courses, describe their planned availability.

Fall
BIOSTATS 690MS Stochastic Models for Population Genomics
BIOSTATS 740 Mixed models and analysis of longitudinal data
EHS 600 Molecular Epidemiology

Spring
BIOSTATS 690T Applied Statistical Genetics
BIOSTATS 690JQ Applied Statistical Learning
BIOSTATS 697G Bayesian Computation in Biostatistics
BIOSTATS 690DS Data Science: Visualization, Statistics and Computing (new course, 2016)
EPI 737 Intermediate Epidemiological Methods
CMPSCI 589 Machine Learning

If the Certificate requires or includes courses from outside the sponsoring department, provide evidence of agreement(s) with the unit(s) offering those courses. You may attach any memoranda of understanding below.

Yes, the certificate involves courses from Computer Science and Environmental Health Sciences. Please see attached letters of support. (Please refer to Proposal #2319 in the Course and Curriculum Management System.)

If the requirements for this Certificate overlap with those of another certificate or a degree or a degree program, describe that overlap. (Note that if a student who has completed a certificate seeks clearance for a degree program that overlaps with that certificate program, the Registrar will note on the transcript that the certificate has been superseded by the degree.)

The Masters in Biostatistics overlaps with the certificate. The certificate could potentially fulfill 4 out of 15 3-credit course requirements for the M.S. in Biostatistics.

What type of student is allowed to participate in this certificate program? (e.g., matriculated UMass students, non-matriculated CPE students, Five College students, graduate students, students in a specific degree program, etc.)

Students who are allowed to participate include qualified undergraduate, graduate and non-matriculated CPE students at UMass.

What role will this Certificate play in relation to other departments or degree programs on campus? Certificates vary widely across campus and may represent a subset of an existing degree program, a multidisciplinary program, or an entirely free-standing area of focus.

This certificate combines the two fields of biostatistics and epidemiology with a focus on application in genomics, and it is thus a multidisciplinary program.
Is this a transitional certificate program?
(Transitional certificate programs are comprised of core courses from specific degree programs and may act as stepping stones into those programs. If a student who has completed a transitional certificate matriculates to the University and completes the degree program associated with that certificate, the transcript will note that the certificate has been superseded by the degree.)

No, this certificate is not intended to be transitional for a Masters Degree in Biostatistics.

If applicable, please attach any additional material relating to the certificate (such as requirement checklists for students, etc.) below.

None at this time.