UNIVERSITY OF MARY WASHINGTON – PROGRAM CHANGE PROPOSAL

Electronically submit this completed form with attachments to the Chair of the College Curriculum Committee.

<table>
<thead>
<tr>
<th>COLLEGE (check one):</th>
<th>Arts and Sciences</th>
<th>x</th>
<th>Business</th>
<th>x</th>
<th>Education</th>
</tr>
</thead>
</table>

Proposal Submitted By: Keith Mellinger  
Date Prepared: 10/23/12

Department /Program:  
CAS Depts of Mathematics and Computer Science; COB

Note: for any program change entailing the addition any new courses, or revisions to existing courses, separate proposal for those course actions must also be submitted.

PROPOSAL TO CHANGE EXISTING PROGRAM (check one of the following)

Revise requirements for existing major  
Revise requirements for a concentration within an existing major  
Revise requirements for an existing degree program  
Revise requirements for existing certificate program  
Revise requirements for existing minor  

Implementation Date: FALL semester, year:

REQUIRED ATTACHMENTS FOR CHANGES TO EXISTING PROGRAMS:
1. Rationale statement (Why is this program change needed? What purposes will it serve?)
2. Impact Statement (Provide details about the Library, space, budget, technology, and impacts created by this program change. Supporting statements from the Library, IT Department, etc. evaluating the resource impact and feasibility of the program change are required.)
3. Catalog Copy (Provide the existing Catalog Description and the complete statement of the proposed new Catalog description that reflects the program changes)

PROPOSAL TO CREATE PROGRAM NOT REQUIRING STATE ACTION (check one of the following)

New concentration within existing major  
New minor  
New Major but NOT a new degree*  

Name: Data Science

Implementation Date (semester and year):

REQUIRED ATTACHMENTS FOR NEW PROGRAMS NOT REQUIRING STATE APPROVAL:
1. Rationale statement (Why is this additional program needed? What purposes will it serve?)
2. Impact Statement (Provide details about the Library, space, budget, technology, and impacts created by this program change. Supporting statements from the Library, IT Department, etc. evaluating the resource impact and feasibility of adding the new program are required.)
3. Catalog Copy (Provide the complete Catalog Description for the proposed new program)

Department Chair Approval: Jennifer Polack-Wahl, Keith Mellinger, Kenneth D. Machande  
Date: 10/31/12

CCC Chair Approval: Bradley Hansen  
Date: 11/15/12

Dean Approval: Richard Finkelstein  
Date: 12/1/12

UCC Chair Approval:  
Date: 

*Provost Approval:  
Date: 

*Required only in cases of proposals for new concentrations, new minors, or new majors that do not involve a new degree
To: Brad Hansen, Chair, CAS Curriculum Committee  
Gail Brooks, Chair, COB Curriculum Committee

From: Keith Mellinger, Chair, Department of Mathematics, on behalf of:  
Karen Anewalt, Department of Computer Science  
Gail Brooks, College of Business  
Stephen Davies, Department of Computer Science  
Julius Esunge, Department of Mathematics  
Chris Garcia, College of Business  
Jangwoon Lee, Department of Mathematics  
Dave Toth, Department of Computer Science

Date: October 23, 2012

RE: Proposal for an interdisciplinary minor in Data Science

The Departments of Mathematics and Computer Science in CAS, together with the College of Business, propose a new interdisciplinary minor in Data Science. The program will be nominally hosted by the Department of Computer Science, but we emphasize that this is purely an administrative designation: all three departments are enthusiastic and equal partners in this program.

The requirements for the Data Science minor are as follows: (Note that the prefix “ANLY,” standing for “Analytics,” is new to UMW. It represents a distinct discipline of study that we hope to build a fifth-year masters program around in coming years. The Data Science minor itself will be a prereq of sorts for entering this Analytics masters program.)

Required courses:

- MATH 200 - Introduction to Statistics (3)
- MATH 300 - Linear Algebra (3)
- CPSC 220 - Computer Science I (4)
- One of:
  - CPSC 230 - Data Structures (4)
  - New course: ANLY 400 - Analytics Application Development (4) (prereq: CPSC 220) - A course in programming and data manipulation techniques for constructing analytics-based applications. Topics include SQL and no-SQL databases, using web service API’s to acquire data, introduction to Hadoop and MapReduce, and use of third-party analytic component API’s. (See separate proposal for this course.)
- CPSC 420 - Modeling & Simulation (3) (to be cross-listed as ANLY 420) (prereq: CPSC 220) -- A robust introduction to techniques of mathematical modeling and computational simulation applied to practical problems. Topics include system dynamics approaches,
discrete–event simulation, cellular automata, and agent–based models. Students complete small projects on topics as diverse as population growth, epidemic transmission, queuing theory, and forest–fire outbreaks. (See separate proposal for this course. This is a significantly revamped version of the older CPSC 420 – “Simulation Techniques," which has not been taught in many years.)

- CPSC 419 – Data Mining (to be cross–listed as ANLY 419) (3)
- One of: (3)
  - CPSC 425 – Parallel Processing
  - New course: ANLY 403 - (prereq: CPSC 220) – Foundations and Applications of Data Analytics (to be cross–listed as BUAD 403) – This course will introduce students to the challenges of developing and applying analytics for insight and decision–making in the real world. Examples and cases will come from customer relation management, price modeling, social media analytics, location analysis, and other interesting areas. (See separate proposal for this course.)

Total: 23 credit hours

**Rationale for why the Data Science minor is needed at UMW.** The pervasiveness of data in the modern digital world is rapidly and fundamentally changing the way organizations operate. Decision making across all levels is increasingly shifting away from subjective human judgment and expert opinion, and is being replaced by superior evidence-based approaches driven by data and analytical models. **Data Science** is the discipline concerned with developing and applying analytical models and methods to gain critical insights from data, understand the behavior of complex systems, and make non-trivial decisions optimally, often in response to quickly changing conditions. Today businesses are employing data scientists to help predict profitable and unprofitable customers, to develop sophisticated product recommendation engines, to determine which patients are likely to be readmitted to a hospital, to adjust hotel prices in real time to maximize revenues, and to inform many other decisions. Natural scientists use the techniques of this field to perform computational simulations in a multitude of areas where actual experiments are impractical or impossible -- such as predicting disease outbreaks, hurricane trajectories, biochemical processes, or the ecological effects of land development.

There is currently a vast shortage of analytical talent in the US. According to a recent McKinsey Institute study[1], there will be a shortage of nearly 200,000 people with deep analytical expertise and another 1.5 million managers who have basic analytics skills by 2018. The *Chronicle of Higher Education* recently identified data-driven computational science as one of “five emerging areas of study,”[2] and it is now widely viewed as the third basic methodology in scientific inquiry, complementing theory and experiment. Accordingly, job demand and beginning salaries are excellent for those with appropriate credentials[3]. In our geographic region in particular, skills like these are highly sought after, as DoD installations and defense contractors search for employees that can bring principled mathematics and efficient computation to bear on weapons and defense projects. In our conversations this summer with representatives from Dahlgren’s Naval Surface Warfare Center, we met with unanimous enthusiasm for developing such a program, and were assured that these skills would greatly strengthen the qualifications of our graduates.
Despite the extensive long-term demand for Data Science, there are currently few programs designed to provide this kind of education. At UMW we are particularly well-poised to address this educational need with faculty from multiple departments and colleges engaged in Data Science research who are making significant contributions to the field. Data Science involves an integrated use of advanced approaches from many disciplines including machine learning, decision analysis, data mining, programming, mathematical modeling, optimization, statistics, and databases. In our business, computer science, and mathematics departments, we have expertise in all these areas, and faculty eager to cross traditional disciplinary lines to collaborate on building a truly cutting-edge program.

The minor itself would serve two complementary functions. First, it would be a natural enhancement to an undergraduate major in mathematics, computer science, economics, business, or the natural sciences. It has great value in that the computational and mathematical skills it delivers are immediately applicable in today’s marketplace. Second, it would serve as the prerequisite for students seeking to complete a masters degree in Analytics, Informatics, or Computational Science, at schools like Northwestern, Indiana University, or perhaps even at UMW in the future. We have selected the coursework required for the minor specifically to instill the background knowledge required for masters work in these areas, since they all require a good deal of skill-building. The creation of this minor represents an attempt for UMW to be ahead of the curve in this contemporary and emerging area of applied science. Such a program would provide a distinctiveness to our already strong academics.

**A minor, not a major.** After lengthy discussions this summer and fall, we have come to the conclusion that the best place for Data Science in the UMW curriculum is as a minor complementing an existing major, rather than a major in itself. The main reason for this is that the skills gained can best be seen as an enabler to applications of another domain. Natural science students, for instance, learn how computational simulations can be brought to bear on the scientific problems they already have the background to understand. Business students carry out data analysis in order to inform the policy questions they have already confronted. Computer science students, proficient in systems programming, learn how to refine these techniques in light of the special implementation challenges that "Big Data" involves. Mathematics students, having studied modeling from a theoretical perspective, then employ Data Science as a practical solution to intractable problems. Rather than "competing" with existing majors, we want Data Science to enhance and expand the learning experience of the students in the excellent majors we already have.

**Comment on course inclusions.** The courses in the minor have been carefully selected with the following criteria in mind: (1) skills necessary to successfully engage in data-science-related projects post-graduation; (2) background necessary to qualify for entrance into masters programs in Analytics, Informatics, or Computational Science; (3) feasibility of completing the minor for our fairly expansive target audience, which includes math, computer science, business, natural science, and social science majors. Data Science is of course a technical discipline that requires a good deal of competence in a broad variety of areas, and it is a challenge to incorporate all these into a minor program. We believe we have done so, however, and that students from all segments of our target audience will find that the minor, though challenging, is not too intimidating to pursue.

Observe that the minor has two places where students can select one of two courses: either CPSC 230 or ANLY 400, and either CPSC 425 or ANLY 403. This reflects a choice in focus: students choosing the first course in these two options convey an interest in the implementation of Data Science programs and systems, and gain more expertise in the practical considerations of their design. Students choosing the second courses, by contrast, focus more on application areas, and how Data Science tools can be
applied to solve problems. The distinction is subtle, but we believe an important one. We expect that computer science majors will normally choose the first option, and business majors the second, but this is by no means a requirement.

Resource impact.

Library. Many of the individual fields in this interdisciplinary venture already have an established presence in our collections, for instance: data mining, mathematical modeling and simulation, statistical methods, high-performance computing, and operations research. It is likely that we would want to add new titles specifically addressing the intersection of the different subfields. Additionally, subscriptions to key journals like *Computational Statistics and Data Analysis*, *Data Mining and Knowledge Discovery*, and the *Journal of Computational Science* would be welcome.

Staff. Faculty within the colleges already have the expertise to teach the desired courses. Some resources will need to be moved around and the College of Business anticipates the need for approximately 1-2 additional adjunct sections per year. The Department of Mathematics believes that their relevant courses can be staffed with only minor strains on resources; a reasonable estimate is that one additional adjunct section may be needed if the minor pushes up enrollments in the minor courses. The Department of Computer Science has seen steady growth in its number of majors in recent years. Anticipating an already increasing demand for seats in CPSC 220 and CPSC 230, the department anticipates the need for three additional instructional adjuncts per year to staff the additional sections required for students seeking the Data Science minor. We have spoken to Dean Finkelstein about this, and though staffing promises can of course not be made, we were assured that this need could very likely be met and would receive the very highest funding priority.

Equipment. Very little equipment is required to support this minor, since most of it involves personal computers already available to faculty and students. Dave Toth from Computer Science is in possession of some significant multi-processor systems for use in parallel processing and high-performance computing, and we anticipate these to be sufficient for that coursework.

Space. We do not anticipate any impact on space resources. The Mathematics and Computer Science departments are already co-located, facilitating collaboration, and when the Convergence Center goes online, we anticipate taking advantage of that space for collaboration between faculty and students from all three departments.

Academic Catalog copy.

Requirements for Data Science minor

Twenty-three (23) credits to include MATH 200; MATH 300; CPSC 220; CPSC 419; CPSC 420; 4 elective credits from among CPSC 230 or ANLY 400; 3 elective credits from among ANLY 403 or CPSC 425. Note that MATH 121 and 122 are prerequisites to MATH 300, and either MATH 201 or CPSC 125 is a prerequisite to MATH 300. Students should bear this in mind when planning their academic coursework.
[3] 2011 graduates of North Carolina State University’s M.S. in Analytics program had an average starting salary of $100,100. The program has held > 90% job placement rate for all four years since its inception.