To the Committee on Honors and Interdisciplinary Studies

I write this letter as a description of, and rationale for, [Name]'s self-designed Data science major. Data Science combines a depth of knowledge on applied design, interdisciplinary critical thinking, visualization and communications with a breadth of knowledge in computer science and mathematics. It can be taught for both generalist students as a single major, or paired with other disciplines (i.e., government, history, international relations, biology, environmental science) as a double major. Data Science distinguishes itself from other fields of study through a focus on communicating meaningful information to non-experts using large datasets; creative applications of existing algorithms (rather than the derivation or proof of new algorithms); and focus on applications in interdisciplinary spaces.

In contrast to “pure” mathematicians, statisticians, or computer and information scientists, a data scientist has a breadth of experience across all of these fields, but may not have as much knowledge as a specialist in any particular field. Critically, a data scientist has a depth of applied experience within at least one knowledge domain, and an exposure to methods for communicating analytic results. This combination allows data scientists to (a) efficiently conduct computational analyses within his/her own knowledge domain coupled with the ability to communicate the results, and (b) manage teams of more specialized individuals to answer far-ranging questions. Individuals with this set of knowledge are revolutionizing a wide set of domains, and are in very high demand not just by faculty researchers at William and Mary, but also by the public and private sector.

[Name]'s self-designed major is designed with all of these principles in mind. We have identified (a) a core sequence consisting of three math courses and four computer science courses; (b) a specialization sequence of two courses which focus on the applications of spatial data in environmental science (GIS and Remote Sensing); (c) a course which focuses on communication (Data 305), and (d) a Capstone COLL 400 experience which will focus on the applications of the skills [Name] learns. This coursework will be further augmented by courses - such as our introductory data science course (COLL 100 - Breaking Intuition) that [Name] selected to take, but we do not formally credit within this major.

The coursework proposed in this major is not offered by any existing program on campus, and thus necessitates [Name]'s request for a self-designed major. Both CS and Math degrees focus on quantitative and technical skills within their discipline, with limited to no focus on applications or communication. The CAMS Applied Statistics program lacks a focus on (a) large datasets, (b) communication, and (c) interdisciplinary applied problem solving, instead focusing on providing students with an (equally important, but different) depth of technical skills. While we used much of this coursework (from Math, CS, and CAMS degrees) as starting points and inputs into the degree [Name] has designed here, it would not be possible to replicate the breadth and applications-focused nature of the degree within existing mechanisms.

Thank you for your consideration of [Name]'s major, and please do not hesitate to reach out with any questions,

Dan Runfola, Ph.D., Institute for the Theory and Practice of International Relations, AidData