Vol 4, No 1, July 2018, 23-36

RESEARCH REPORT

Cross-Disciplinary Studies Minors as a New Vehicle to Enhance STEAM Programs

Alex Dekhtyar [a], Andrew Schaffner

California Polytechnic State University, USA

Abstract: *This paper describes the motivation and organization of Cal Poly's new type of academic subprogram: the Cross-Disciplinary Studies Minor (CDSM). CDSMs are an appropriate educational vehicle for providing both depth in one's own field and breadth in a companion field for students who want to excel in studying topics that straddle the boundaries of traditional BS degrees. The design of CDSMs is illustrated using the example of Cal Poly's Cross-Disciplinary Studies Minor in Data Science, which has been in place since 2015-2016 academic year.*

Keywords: STEAM, Cross-Disciplinary Studies, Cal Poly

Introduction

In September 2015, two brand-new programs opened their doors at Cal Poly, San Luis Obispo: the Cross-Disciplinary Studies Minor in Data Science[1] and the Cross-Disciplinary Studies Minor in Computational Interactive Art[2]. The common denominator connecting these nascent programs is a new type of academic subprogram: The Cross-Disciplinary Studies Minor (CDSM). We designed it to specifically address a number of challenges that traditional academic programs and subprograms face when dealing with cross-disciplinary education. In particular, the CDSM provides an academic policy to structure and coordinate an in-depth curriculum in partnered related disciplines with the aim of satisfying the changing needs of industry for employees to have strong skill sets integrated across multiple domains. The introduction of CDSMs at Cal Poly led to immediate development of the current programs in Data Science and Computational Interactive Arts. Another program, the Cross-Disciplinary Studies Minor in Bioinformatics, is currently under development for anticipated enrollment in Fall 2019.

In this paper, we describe the motivation, challenges, and process behind developing Cross-Disciplinary Studies Minors and describe their structure. Using Cal Poly's CDSM in Data Science as a case study for implementation, we show how specific curricula for CDSMs can be developed. We concentrate our discussion on how the Data Science CDSM addresses the challenges faced by Cal Poly, while leveraging the strengths of existing BS programs in statistics and computer science, and on how this experience can be replicated by other institutions of higher learning to spearhead quality cross-disciplinary educational efforts.

The rest of this paper is organized as follows. We first discuss the emerging need for cross-disciplinary education programs at institutions of higher learning, and outline the challenges to successful design and implementation of such programs. We then introduce the key part of our solution: the Cross-Disciplinary Studies Minors. After that we discuss our implementation case study, the curriculum for the CDSM in Data Science. We conclude with a few observations and lessons learned.

[a] Department of Computer Science and Software Engineering, Department of Statistics California Polytechnic State University, San Luis Obispo, E-mail: dekhtyar@calpoly.edu

Dekhtyar, A., & Schaffner, A. (2018). Cross-disciplinary studies minors as a new vehicle to enhance STEAM programs. *Journal of Research in STEM Education*, *4*(1), 23-36.

[1]http://catalog.calpoly.edu/collegesandprograms/collegeofsciencemathematics/statistics/ crossdisciplinarystudiesminordatascience
[2]http://catalog.calpoly.edu/collegesandprograms/collegeofengineering/computersciencesoftwareengineering/computingforinteractiveartsminor

ISSN:2149-8504 (online)

Motivation

Background

California Polytechnic State University (a.k.a. Cal Poly, San Luis Obispo) is a comprehensive public university of about 20,000 students and is one of the 23 campuses of the California State University system. Cal Poly is classified as a regional Masters-level institution and is consistently ranked as the best public university in the West in its category[3]. Cal Poly consists of six academic colleges operating on the quarter system. Our institutional motto is "learn by doing," and this motto is incorporated at all levels of curricula across campus, promoting hands-on, laboratory-based and project-based learning experiences.

The curriculum development process at Cal Poly traditionally originates within a single department and subsequent review running through the department's college followed by the Academic Senate Curriculum Committee. Final recommendation for approval is granted by the Academic Senate with final approval granted by the Provost. Prior to 2015, when Cross-Disciplinary Studies Minors were introduced, Cal Poly had a limited track record of development and support of cross-disciplinary programs. In fact, some past administrative actions actually made cooperation among related disciplines more difficult. For example, at the undergraduate level, the establishment of a BS degree in Computer Engineering in 1984 lead to a split of the Department of Statistics and Computer Science into its constituent parts, with the subsequent move of the Department of Computer Science from the College of Science and Mathematics to the College of Engineering, where the Department of Electrical Engineering, the other collaborator in the Computer Engineering program resided. Other administrative practices only carry the illusion of being cross-disciplinary. Consider the graduate level cross-college dual-degree offered at Cal Poly: a combined MBA and MS in Engineering Management. This dual-degree program essentially pairs two existing 45-unit[4] MS degrees into a single 90-unit course of study without any significant coordination of the partnering degree curriculum.

Emergence of the Need for Cross-Disciplinary Educational Experiences

Today cross-disciplinary knowledge is either outright required or is in high demand for some college graduates. Examples include the applied fields of data science, bioinformatics, and a variety of management and entrepreneurial occupations that must pair business knowledge and experience with a specific domain (Kim & Lee, 2016; Finzer, 2013; Lafferty-White, 2017; Levine, 2014; Davenport, 2012; Dymond et al.; 2015; Luke et al., 2015).

We define a cross-disciplinary program as one that combines in-depth knowledge in two or more related fields with additional advanced study at the intersection of the fields. For example, our data science cross-disciplinary curriculum combines in-depth coursework from the fields of computer science and statistics; our computing and interactive arts cross-disciplinary program – from computer science and art; and currently under development at Cal Poly, bioinformatics – from computer science, statistics, and biology. Other examples of cross-disciplinary studies include health management – from business and pre-med/biology; public health – from social science, nutrition, and biology; and bio-innovation – from business and biology. The term "cross-disciplinary" is chosen to emphasize the need for a single person (a student, in case of an academic cross-disciplinary program) to master the advanced knowledge in multiple disciplines. This is in contrast to other common -disciplinary terms such "multidisciplinary" or "interdisciplinary", which we take to describe team compositions comprised of expert individuals from multiple fields with possibly only rudimentary understanding in partner fields[5] (Knight et al., 2012; Razzaq et al., 2013).

[3]See http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/regional-universities-west. Cal Poly is ranked #11 overall and is the top-ranked public university.

[4]All units mentioned in this paper are quarter units. Thus a 45-unit MS degree at Cal Poly is equivalent to a standard 30-unit degree at university that uses the semester system.

ISSN:2149-8504 (online)

When considering the need for new curricular programs, an important first question is who the students to be trained in the new field are. With cross-disciplinary studies, universities have an existing built-in population of students who can be engaged. These are the students majoring in one of the partnered disciplines associated with the cross-disciplinary study. For example, a target population of students to study bioinformatics can come from Biology, Computer Science, and Statistics majors mutually interested in focused study (beyond a traditional minor) in the complementary disciplines in addition to their own. We note, that this actually leads to a certain diversity of experience and expertise among the students trained in cross-disciplinary studies: a student with a BS in Biology trained in bioinformatics will have more in-depth knowledge of biology and less in-depth knowledge of Computer Science than a CS major trained in bioinformatics. However, both will have a greater depth of knowledge in the partner area compared to those students who simply minor in the partner disciplines. Additionally, unlike those who complete traditional minors, they will have completed advanced studies at the intersection of the disciplines and greater experience working on diverse teams.

Industry wants trained professionals with cross-disciplinary skills; multidisciplinary teams are also important, but not sufficient in many cases. Universities have a limited number of options on how to respond. Some universities can afford to establish new BS programs run by brand new departments with new tenure lines. This is expensive, may create duplication of experience and coursework, and incite turf wars.

Another common option is post-baccalaureate degree education through certificates. As an example, a BS in Biology can take a certificate in CS-centric bioinformatics (Parthasarathy, 2015; Ranganathan, 2005). The drawbacks here stem from the fact that there are essentially two completely independent paths to bioinformatics proficiency - one for those with BS in Biology, and the other - for those with BS in Computer Science (not to mention a third path for BS in Statistics degree holders). These varied paths, and hence student preparedness, limit the depth of coursework in the intersection. Additionally, certificates are constrained by the limited amount of coursework they can have.

A third option is an MS degree in a cross-disciplinary subject. This, for example, is a popular emerging path with Data Science (Tate, 2017). While BS in Data Science degrees are still fairly rare, many universities have created MS in Data Science (or similar) programs hoping to attract applicants with a wide range of undergraduate degrees. A cursory look at these programs though reveals their key disadvantage. While MS in Computer Science and MS in Statistics programs are true in-depth advanced degree programs in their respective fields, MS in Data Science education often starts with remedial subjects, such as Linear Algebra, Numerical Analysis, Introduction of Computing and Introduction to Statistics, in order to accommodate the wide variety of educational backgrounds and skills of the incoming class. This leaves precious little time for actual in-depth preparation in the core field of study.

Impediments to Cross-Disciplinary Learning Using Traditional Undergraduate Programs

There are several core obstacles preventing successful deployment of full-scale "ideal world" cross-disciplinary studies curricula at modern US universities. In particular, they are fueled by the non-trivial desiderata for such curricula.

In-depth education in multiple fields of study. Industry needs are not properly addressed by BS degree holders with only rudimentary command of one of the partner disciplines (Pournaras, 2017; Kim and Lee, 2016; Van der Alst; 2014). A college graduate hoping to continue their career in the field of bioinformatics would need to have in-depth knowledge of biology – perhaps not of the entire field, but certainly of cell and molecular biology, genetics and evolution. The same graduate needs to be a competent software developer, with CS expertise that is above and beyond the introductory CS course sequence (e.g., databases, algorithms,

[5] We understand that the use of "-disciplinary" terms such as "interdisciplinary" or "cross-disciplinary" in other work may differ from ours, since these terms are not fully codified. Rather than attempting to instill our readings as general definitions, we simply clarify the use of these terms in this paper. When referring to prior work, we use the term (usually "interdisciplinary") used in the work itself, but elsewhere, we stick to our interpretation of "cross-disciplinary" vs. other similar terms.

ISSN:2149-8504 (online)

machine learning techniques, advanced data structures, etc.). This need for depth in multiple disciplines is a challenge because at present, most existing BS programs provide for in-depth study in only a single field.

Budget. Ideally, a university can hire qualified faculty in a new emerging cross-disciplinary area, create a new department (if necessary), and a create a new program of study from scratch. Realistically, however, universities are guided by the same return-on-investment principles that guide business decision-making. A cross-disciplinary program has a higher chance of being approved by administration and a higher chance of having institutional support if it relies on existing resources and strengths, and requires only modest additional investment from the administration. This typically means, that new BS degree programs are often not achievable in the contexts of US universities. This is doubly so for resource-constrained universities, such as those with a focus on teaching and those without Ph.D. programs.

Filling the Gaps. There is a modest scope of literature that defines interdisciplinary studies as a means to explore these gaps or intersections. An entire journal dedicated to this topic: Issues in Interdisciplinary Studies. We note, however that much of the interdisciplinary literature emphasizes partnerships among experts from multiple disciplines to solve interdisciplinary problems as opposed to experts in the intersection between disciplines. "Interdisciplinary research is pluralistic in its methods and involves researchers working in tandem with each other in an integrated way to create new and unpredictable patterns, referred to as a 'kaleidoscope'" (Razzaq et al. (2013)). In contrast trans- or cross-disciplinary research refers to the intersection of disciplines to be a new discipline in its own right: "collaboration and mutual learning among people from practice and society are a salient and necessary part of transdisciplinarity." (Razzaq et al. (2013)).

While there is often confusion, misuse, and exchangeability using the terms inter, multi, cross, and trans-disciplinary as synonymous gradations, we aim to directly train students towards becoming an expert in the new intersection domains – the gap.

Augsburg and Henry (2009) and Klein (2009) define the notions of strong and weak interdisciplinary programs touching on one of the most important issues with building cross-disciplinary studies programs: they characterize strong programs by the existence of special-purpose coursework that "bridges the gap" between the partnered fields of study, while weak interdisciplinary programs are characterized by the absence of such coursework. It is insufficient for a cross-disciplinary study program to merely bring together the knowledge from two related fields by offering curriculum in only the partnered fields. Coursework to facilitate the synthesis of knowledge and skills from both fields, not taught in existing coursework in either field, must be provided. This synthesis coursework represents "the gap" between the two disciplines that a Cross-Disciplinary studies program needs to bridge. As examples of coursework indicative of a strong program, Augsburg and Henry (2009) identify (1) the presence of an introduction into the cross-disciplinary area courses (such as Introduction to Data Science for a Data Science program), and (2) capstone coursework that synthesizes the knowledge and applies it to specific project-oriented work. The ability to offer such book-end coursework is essential; it strengthens the students' command of the respective subject matter, allows them hands-on experience within the actual cross-disciplinary field, and teaches them how to apply discipline-specific knowledge and skills in problem-solving in the cross-disciplinary field. These "bridging the gap" synthesis classes would typically only be accessible to the students pursuing the cross-disciplinary studies, as they would require drawing heavily on the prior in-depth multi-field knowledge that only those students gain.

Integration of cross-disciplinary study throughout the entire program. Many possible ways of offering cross-disciplinary educational experiences assume that students already have received a degree, or have been exposed in-depth to one of the fields related to the cross-disciplinary study, and thus concentrate on remediating their lack of knowledge and skills in the other related field or fields. In contrast, a well-designed cross-disciplinary study program can provide the exposure to the coursework in all related fields in an integrated way. For example, statistics students studying Data Science progress through the introductory coursework in both Statistics and Computer Science during their first two years in college, rather than learning how to proficiently program in Python after they have received a BS in Statistics (as an example). Similarly, because of their early experience

ISSN:2149-8504 (online)

in multiple fields, during their third and fourth years of study, both computer science and statistics students in data science can be exposed to advanced coursework relevant to data science (e.g., Regression Analysis from the Statistics Curriculum, and Distributed Computing from the CS curriculum) in parallel. This creates a balanced "leveling-up" of students, and allows for appropriate "bridge-the-gap" coursework to complete the program of study. Further, because students begin taking partner work early in their academic year, students will be able to achieve the cross disciplinary objectives and graduate in a timely manner.

Limitations of traditional programs and subprograms

Cal Poly offers two types of primary degree experiences: an undergraduate (B.S.) degree and a graduate (M.S.) degree. Prior to the development of the CDSM, Cal Poly offered three non-degree granting subprograms: undergraduate minors and concentrations, and graduate specializations.

Budget constraints make it difficult to offer cross-disciplinary experiences such as Data Science, Computational Interactive Art, or Bioinformatics as independent BS or MS degree programs. Cross-disciplinary education delivered in a form of subprogram is potentially more flexible, nimble, and cost-effective in the context of a school like Cal Poly.

As noted above, we argue that three important components comprise the cross disciplinary curriculum: (1) an introduction to the cross-disciplinary study, (2) in-depth coursework in each of the partner domains, and (3) synergistic courses and/or capstone experience. Minors and concentrations cannot meet these demands with limited units (a maximum of 30 additional quarter-units for minors and a maximum of 180 units for BS degrees with a concentration). Because of minor or degree unit limits, traditional subprograms either provide (1) a lack of depth in the partner discipline (minors), or (2) a lack of exposure to the partner discipline altogether (concentrations).

Unit limits aren't the only inadequacies of concentrations and minors for cross-disciplinary studies. Because concentrations are limited to subprograms within a single major, it is not possible to maintain a common curriculum for all students in a cross-disciplinary study across multiple (independently managed) BS degrees. And, because a core principle of most minors is exposure and breadth with an "everyone is welcome" philosophy, minor programs lack the needed depth to prepare students for the advanced synergistic capstone coursework in the cross-disciplinary study.

A new subprogram: the cross-disciplinary studies minor

To administratively address the requirements for cross-disciplinary studies we developed the cross-disciplinary studies minor. The new subprogram was adopted by the Cal Poly Academic Senate and approved by the University President in the 2013-2014 academic year as permitted by CSU Executive Order 1071, which grants each CSU campus the responsibility for defining the rules and restrictions for subprograms. While many universities have provisions for students to individually create programs of interdisciplinary studies, to our knowledge, the structure, goals, and requirements of our CDSM subprogram is unique among US academic institutions as a formal subprogram to address cross-disciplinary domains. The definition and key requirements of the CDSM as adopted at Cal Poly are as follows.

Definition. A cross-disciplinary studies minor (CDSM) is the result of a partnership between two or more target major programs. It is defined as a set of curricular requirements comprised of coherent groups of courses tailored for each partner program such that all students from target majors develop (1) depth in the partner discipline, (2) focused study in their own discipline, as well as (3) focused study in the mutual domain of the minor.

[5] Now officially known as the Department of Computer Science and Software Engineering.

ISSN:2149-8504 (online)

Key Requirements

1. The curricular requirements are the same for all students in the CDSM.
2. The total number of units in the CDSM that cannot be covered by the requirements of the student's major shall not exceed 24 quarter-units.
3. The CDSM curriculum shall require at least 12 quarter-units of coursework that cannot be covered by the requirements of the student's major.
4. At least half of the units must be from upper division courses.

Each key requirement is readily justified.

Identical curricular requirements regardless of student's major. This follows from the principle that all students pursuing a CDSM, regardless of their major are to acquire the same knowledge and skills. The intent is for the CDSM creators to identify the list of courses in both partner disciplines which provide the appropriate background, skills, and knowledge, and require these classes of all students. The additional new coursework developed specifically for the minor is also required for all students. At the same time, while students from all target majors have the same curricular requirements, some requirements may be met by more than one course.

Limited total units. As an undergraduate degree enhancement, only a reasonable number of units may be required without fully impeding a timely graduation or turning the program into a full dual-degree. Cal Poly's traditional minor policy requires between 24 and 30 units of coursework.

Minimum required units beyond the major. As a cross-disciplinary study, the coursework should not be entirely subsumed by courses already in the major BS curriculum. As a degree enhancement, we require that students take coursework beyond the minimum requirements for the BS degree.

Upper division coursework. Similar to the traditional minor, advanced coursework is required to ensure discipline depth. In the traditional minor this coursework is often chosen from a menu of courses, whereas the cross-disciplinary minor will usually explicitly require certain upper division courses that support the cross-discipline.

Steps for developing partnerships and curriculum for a CSDM

In the development of curriculum for new CDSM we suggest following the following steps: (1) identify stakeholders, (2) generate learning objectives, (3) identify existing curriculum and curricular gaps to support the learning objectives, and (4) streamline and map the curriculum to partner programs.

Identify stakeholders. There are three primary stakeholder groups to consider when developing the new CDSM: students, faculty, and industry. The natural student and faculty populations stem from the partner disciplines in the cross-disciplinary studies. We recommend limiting the number of partner programs in order to preserve the ability to maintain sufficient depth in each of the partner fields and to better facilitate curricular maintenance as programs evolve.

Generate learning objectives based on understanding of the cross-disciplinary needs and industry needs. The essential questions to ask at this stage are "what do we want our students to learn in the CDSM?" and "what do we want them to be able to do upon completion?" Understanding the demands of today's workplace, the work the professionals who complete the CDSM will be asked to perform, can help formulate the answers to these questions. We developed CDSMs to produce career-ready employable graduates with practical experience. CDSM students will also be well prepared to pursue graduate coursework, but the curriculum emphasizes career readiness within the curricular limits of an undergraduate program.

Identify existing curriculum and curricular gaps. The next step is to map the high-level learning objectives produced in the previous stage to specific knowledge and skills that are to be taught in the CDSM. In this step, each item from the list of objectives is mapped to courses in the curricula of the partner disciplines in

ISSN:2149-8504 (online)

which it is taught, or is labeled as "unmet" by the existing curricula. Any learning objectives, skills, or knowledge that are not sufficiently addressed with existing curricula will highlight gap needs that can be addressed both an in introductory course in the CDSM, as well as in the synthesis coursework.

Streamline and map curriculum to partner programs. The reason why CDSMs are useful is their unique ability to map courses required by the CDSM to the curricula requirements of the partner programs in ways that drastically reduce the course load above and beyond the major requirements. This needs to be done separately for each partner discipline. Because close to half (or, perhaps, close to 40%) of the CDSM may consist of courses coming from a specific partner discipline, students majoring in this discipline will have some of the courses already required in their major program while other courses may count as electives. Sometimes, it may also be possible for students to count the coursework offered by the other partner program as electives in their major: for example the CDSM in Data Science (see below) takes full advantage of the flexibility of both the CS and Statistics curricula to incorporate 4-5 courses from the partner discipline as electives of different kinds.

The process of mapping the courses in the CDSM to each partner discipline, while pruning the list of courses as needed to be able to meet the "no more than 24 units above and beyond the major" requirement is, in our experience, the most complex part of the CDSM development process. The initial list of potential coursework to address the objectives from each of the partners can be quite large and must be pruned. In the mapping and pruning process we strive to balance coursework from each of the partners, as the resulting pruned curriculum will have to be mapped to each partner BS degree program. We shape the curriculum to have an equitable balance of extra-discipline (beyond credit for the major) coursework for each partner program as well as equitable unit double-counting for degree progress. In some cases, in order to be able to map courses in CDSM to the BS degree requirements in desired ways, changes to BS degree requirements, such as adding certain courses from one discipline to the list of electives of another discipline, or changing course prerequisites, may be proposed. In such cases, CDSM designers have to confer with the appropriate curriculum committees to ensure that the proposed changes will not have adverse effect on the BS degree itself.

Our Data Science case study demonstrates how the final list courses for the CDSM was selected and mapped to the BS in CS and BS in Statistics requirements, and how the CS and Statistics curricula was updated to accommodate the CDSM.

Case Study: Cross-Disciplinary Studies Minor in Data Science

The notion of Cross-Disciplinary Studies Minors appeared during the two-year long process of developing a Data Science program at Cal Poly undertaken by the authors of this paper. In this Section we briefly introduce the stakeholders in the minor, give the short history of its development, and discuss the CDSM in Data Science as currently deployed at Cal Poly, paying close attention to the aspects of the minor discussed in the previous section.

The Stakeholders

Early in the development of the Data Science program (and before it became the CDSM in Data Science), the Departments of Computer Science[6] and Statistics emerged as two key stakeholders in bringing Data Science education to Cal Poly.

The Department of Computer Science oversees the B.S. in Computer Science, B.S. in Software Engineering, M.S. in Computer Science, Computer Science Minor program, and jointly with the Department of Electrical Engineering houses the B.S. in Computer Engineering program. Between the three majors and the M.S. program (and not including the hard-to-count students minoring in CS) the department serves around 1200-1300 students every given year.

The Department of Statistics runs a B.S. in Statistics program with the total enrollment of about 120-140 students as well as a Statistics minor serving approximately 90-100 students. It also hosts an Actuarial Sciences Minor. In addition to teaching its own majors, the department offers a wide range of Statistics coursework to

ISSN:2149-8504 (online)

students from other disciplines: the vast majority of Cal Poly majors require at least one course in Statistics.

Brief History

In Spring of 2012, a conversation around the subject of Data Science took place on the Cal Poly campus, featuring interested parties from all colleges, and from many departments. This conversation gave birth to numerous initiatives, chief among which was the emergence and approval of the Cross-Disciplinary Studies Minors at Cal Poly, and the establishment of the CDSM in Data Science.

As part of the initial conversation about the Data Science education at Cal Poly, we asked the question of who, among the current body of our students, are most ready for Data Science education. After some conversation about the definition of Data Science, and the skills and knowledge associated with professionals trained in this area it became clear that two existing student populations: Computer Science majors and Statistics majors already take much of the coursework that was deemed necessary for a proper Data Science education.

A preliminary survey conducted in 2012-2013 suggested that around 35% of Statistics majors, and 13% of CS majors may be interested in a Data Science curriculum. Taking the respective sizes of the programs into account, we estimated that an undergraduate Data Science program that relies on CS and Statistics students as primary target populations can realistically have 20-30 students per year/cohort roughly evenly split between the two majors. Our program design used this estimate throughout the process, and, at present, it appears that our estimate was correct: the 2017-2018 (our second) cohort is 24 students.

With CS and Statistics majors in mind we approached the development of our Data Science curriculum in the following way. At the outset, we put together an overarching list of skills that we wanted Cal Poly graduates of a Data Science program to have. The list ranged from relatively generic skills, such as "able to work with databases to store, organize, and retrieve data" to some very specific ones, such as "able to apply Principal Components Analysis to reduce the dimensionality of complex data". We then mapped the desired skills to the existing CS and Statistics coursework, tracked all course prerequisites (e.g., databases are covered in Introduction to Databases course, its prerequisite is a Data Structures course), and identified "gap skills", i.e. skills and experiences that the students were not getting from the CS and Statistics courses. We determined that at a bare minimum the "gap skills" could be covered in four courses (totaling 12 quarter units): an Introduction to Data Science course, a senior-level Data Science course, and a two-quarter capstone sequence. The total number of quarter units in the coursework we earmarked for a Data Science curriculum (including the new coursework) was 80 – far exceeding the unit requirements for a minor.

This was discouraging. But we did not stop there.

We asked ourselves, if a B.S. student majoring in Computer Science (Statistics) were to take all of these courses, how many extra units would it cost them? Our mapping of the 80 units of proposed coursework (see below) to the B.S. in CS and B.S. in Statistics degree requirements surprised even us:

- with minor alterations in the Statistics curriculum, a Statistics major could take all the courses with only 12 extra units above and beyond the 180 major units.
- with minor alterations in the Computer Science curriculum, a Computer Science major could take all the courses with only 16 extra units above and beyond the 180 major units.

This feat is possible because much of the coursework in one's own major is either required in one's own degree, or counts as a technical elective, and because both the B.S. in Statistics and B.S. in Computer Science are flexible enough to allow counting upper-division courses taken in the other major as technical electives.

At this point we made use of CSU Executive Order 1071 which grants the responsibility for subprogram development to individual campuses to propose a new subprogram type that would allow us to offer the Data Science curriculum the way we envisioned it without having to compromise. The proposal for Cross-Disciplinary Studies Minors was created in consultation with a number of other possible stakeholders on campus (faculty from departments who could partner for additional CDSMs), and was shepherded through the Academic Senate.

The Curriculum

Matching and enhancing the target majors, the Data Science curriculum consists of coursework in three areas: traditional Statistics (with some prerequisite math), traditional Computer Science, and newly developed Data Science Courses. The courses range from full introductions to both disciplines through upper-division technical electives. All students studying for the Data Science CDSM will take the courses described below, plus two elective courses of their choosing in computer science, statistics, or data science.

Statistics and Math Coursework

Two mathematics courses were included in the official description of the Data Science curriculum: Calculus III (as a required prerequisite to Linear Algebra), and Linear Algebra. Knowledge of Linear Algebra was deemed a core requirement for a properly educated Data Scientist.

The Statistics coursework consists of five courses. Two are introductory in nature: an Introduction to Statistics course, an Introduction to Probability Theory. Three other courses are Applied Linear Regression, Multivariate Regression, and Statistical Computing in R. These classes from the Statistics curriculum address several of the statistical objectives deemed necessary for the Data Science CDSM: linear regression, knowledge of modern tools for statistical analysis (R), and dimensionality reduction techniques using Principal Component Analysis.

Computer Science Coursework

Four courses from the introductory CS curriculum became part of the Data Science curriculum: Fundamentals of Computer Science I, Fundamentals of Computer Science II (Object-Oriented Programming), Fundamentals of Computer Science III (Data Structures), and a Discrete Mathematics course. In addition, four upper-division courses were added to the curriculum to cover the requirements for Data Scientists to be familiar with algorithm development, databases, distributed computing, and machine learning techniques. The four courses are Design and Analysis of Algorithms, Introduction to Databases, Distributed Computing, and Knowledge Discovery in Data (the CS department's undergraduate Data Mining and Machine learning course).

Data Science Coursework

The remaining four required courses in the CDSM are specially developed for it to address gaps in the cross-disciplinary curriculum.

Introduction to Data Science. This sophomore level gateway course is intended to be accessible to students after completing only an introductory statistics course and two-quarters of introductory programming. Primarily using Python, students get their first exposure to working with non-standard data types including raw machine, text, geospatial, images, and audio. The data science workflow is introduced in small scale settings. Based on interest and success in this course, students decide whether or not to continue (or be accepted into) the CDSM.

Data Science. This senior level course (team-taught by CS and Statistics faculty) is taken during the Fall quarter of the senior year after completing the majority of the CS and Statistics courses in the CDSM. This course aims to synthesize the statistics and CS skills acquired in the prerequisite coursework and to present a holistic picture of Machine Learning not seen elsewhere in the curriculum. The course introduces core material from the fields of applied mathematics, statistics, and computer science. Students receive a brief introduction into numerical analysis and optimization, including gradient descent and Lagrangian optimization. Drawing from the statistics discipline, the course presents parametric and nonparametric models of data, and the principle of maximum likelihood to derive a number of machine learning techniques such as linear and logistic regression. Advanced machine learning techniques not covered in the Knowledge Discovery from Data course,

ISSN:2149-8504 (online)

such as Support Vector Machines and Neural Networks, are also studied with emphasis on the mathematics behind the methods and their optimization techniques to discover solutions. Throughout the course students work in small teams on a number of data science projects to practice applying the material covered in the class culminating in class presentations.

Data Science Capstone I and II. This 6-month long capstone sequence follows the Data Science course (above) and completes the senior year. This capstone is meant to give students real-world experience so that they can be ready for industrial and scientific applications. Working with real clients, student teams manage all aspects of the data science workflow. In collaboration with their clients, they specify and design project requirements, implement data gathering methods leading to the deployment and delivery of a system or analytical methodology that involves working with, analyzing, and visualizing large quantities of data. Projects include technical documentation, quality assurance, integration and systems testing. Other non-technical skills include project planning, time and budget estimating, project team organization, ethics, and professionalism. The inaugural capstone was taught in Winter and Spring of 2017 to a cohort of nine students. The course had three outside customers: a government agency (USAID; students working on this project analyzed household surveys from a number of African countries), an academic customer (a Political Science professor interested in the problem of redistricting), and an industry customer (a San Francisco startup OpenDoors, which asked the students to learn a variety of features for evaluating the price of real estate).

Making it Work

The full CDSM in Data Science proposal, as described above constituted 21 required courses totalling 80 quarter units, including four new courses developed specifically for the Data Science curriculum (12 units). This is a very ambitious curriculum: by way of comparison, BS in Computer Science major requires 15 CS courses, seven more technical electives and a two-quarter senior project for a total of 24 courses in the major. The key behind the ability to offer an ambitious cross-disciplinary Data Science curriculum to Computer Science and Statistics majors lies in the ability to map the CDSM coursework to the respective BS requirements in CS and Statistics. Table 1 shows how each course was mapped to both the Computer Science and Statistics curricula.

As seen from Table 1, both mappings take significant advantage of the fact that many courses in the curriculum are required for each respective major, and that there is reasonable flexibility in the elective coursework in both majors. Eight courses from the list above are accounted for via major requirements in one or more programs (five mutually required). Additionally, BS in CS students can count six more courses as various types of electives, including all but one upper-division Statistics courses. BS in Statistics students can count seven courses as various types of electives, including all upper division CS coursework. Only the four special purpose Data Science courses (and the Multivariate Analysis course for CS majors) comprise the coursework above and beyond the major.

ISSN:2149-8504 (online)

Table 1.

Mapping of Data Science curriculum classes to CS and Statistics majors.

Course	BS in CS	CS: extra units	BS in Statistics	Statistics: extra units
Calculus III	required	0	required	0
Linear Algebra	required	0	required	0
Intro to Statistics	required	0	required	0
Intro to Probability	math elective	0	required	0
Applied Linear Regression	upper div elective	0	required	0
Multivariate Analysis	minor	**4**	tech elective	0
Statistical Computing in R	tech elective	0	required	0
Fundamentals of CS I	required	0	required	0
Object-Oriented Programming	required	0	support elective	0
Data Structures	required	0	support elective	0
Discrete Math	required	0	required	0
Algorithms	required	0	tech elective	0
Intro to Databases	tech elective	0	tech elective	0
Distributed Computing	tech elective	0	tech elective	0
Knowledge Discovery in Data	tech elective	0	free elective	0
Intro to Data Science	minor	**4**	minor	**4**
Data Science	minor	**4**	minor	**4**
Data Science Capstone I	minor	**2**	minor	**2**
Data Science Capstone II	minor	**2**	minor	**2**

Lessons Learned

The Data Science CDSM opened its doors in Fall 2015. Due to the structure of the minor, it was possible to engage rising sophomores and juniors, thus allowing us to produce our first graduating class of nine students representing four majors (Statistics, CS, Computer Engineering and Mathematics) in Spring of 2017. Another 24 students will receive the Data Science minor in Spring 2018. It will take several more years for us to be able to appropriately assess the actual impact the program has had on its participants. At the same time, we believe that our experience and discoveries on the way to designing the CDSM may offer valuable insight to other campuses in how to proceed in building a healthy cross-disciplinary programs in ways that leverage the existing knowledge and expertise to a maximum. Throughout our experience, the following core lessons and conclusions have emerged.

When everything else fails, create your own rules. To create a successful data science program, we had to build a brand-new unique academic subprogram: the Cross-Disciplinary Studies Minor. The CDSM is a perfect vehicle for programs that want to combine a concentration in a student's major, with an in-depth minor in another, related and relevant area: the key feature of CDSM is to make such opportunities reciprocal for participating programs. The authors are grateful to all their colleagues among the faculty and administration who have participated in the past five years in discussions and the approval process for this vehicle.

The CDSM is a degree enhancement. The initial development of the Data Science curriculum followed a path similar to the development of a full-blown degree program, but to achieve all of the learning objectives (and general education requirements of all degree programs) the total unit demand would exceed 180 units, the maximum permitted by Cal Poly policy. For example, using the CDSM as a vehicle to offer a curriculum to meet the full set of objectives yields a "super-degree" in Data Science – the BS enhanced by the CDSM. In general, a

ISSN:2149-8504 (online)

key feature of a CDSM is that it produces two (or more) "flavors" of specialists: one for each participating degree program. For example, the CDSM in Data Science produces Data Scientists with BS degrees in Computer Science and Statistics, each with their own, unique to their core discipline additional knowledge and skills.

CDSMs allow for bigger subprograms. The immediate consequence of a definition of a CDSM is that with a pair of participating programs, it provides an immediate ability to build either a 12-course cross-disciplinary curriculum with little additional synthesis coursework, or a 10-course one including more synthesis coursework for universities operating on the quarter system. For universities operating on semester system, a CDSM that does not exceed the units above and beyond those of a regular minor would allow for an 8-course "weak" curriculum, or a 6-course "strong" curriculum.

We illustrate this on the example of the other CDSM at Cal Poly, the Computing for Interactive Arts (CIA) minor. The participating programs in this CDSM are Art and Computer Science. These programs have no ability to count courses taken in the partner discipline towards their own own BS degree. As such, the CIA minor incorporates three required Art courses, three required CS courses, two approved tech electives in each field, and a two quarter capstone course (totaling 4 units) for a total of 12 courses. Each major counts the five courses (20 units) in its field as part of the BS degree, and leaves the remaining seven courses worth 24 units as above and beyond the major. The CIA CDSM does not have an introductory course in the field of Computational Interactive Art, but it does include a two-quarter project-based capstone.

Widespread campus interest. Our development of the CDSM in Data Science was an attempt to directly address the growing demand for formally trained ready-to-deploy data scientists by developing a rigorous curriculum with focus on foundational and generalizable concepts in computer science and statistics. This approach excludes any training for the more casual analyst who might use a commercial visualization tool on a processed or structured dataset. As a result, colleagues in Business and Journalism, for example, have felt alienated and have developed curriculum of their own to address niche analytics (e.g., in Business, an M.S. in Business Analytics; in Journalism a course in "Data Journalism"). We have been challenged to find a "one-size fits all" solution. At the same time, consistent with the core demands in the tech industry for properly trained data scientists, we believe that a curriculum that emphasizes the technical depths in CS and statistics and targets students in these majors needs to be offered. Note that access to the CDSM in Data Science is not exclusive, it is merely optimized for the statistics and computer science programs. Some of the students who have already received the Data Science CDSM, or have been admitted into the minor are students in mathematics, computer engineering, software engineering, and political science majors. Their participation in the program has a higher unit cost, but a cost these individual students are willing to bear.

Collaborative nature of the process. To put it simply, this achievement would not have been possible without two critical components: the ability of the representatives of the two departments to work together in a trusting manner on the development of the CDSM proposal. Additionally, each specific CDSM requires strong support from the participating departments, including the willingness of the departments to admit additional students into their classes, collaboration between the CDSM creators and the department curricula committees and department chairs.

Unique Educators. If it's difficult for industry to find competent young data scientists, it is even harder to find data scientist educators. As a still emerging field, there are still very few new PhDs who are comprehensively trained, and those that are have many opportunities for financial reward that greatly exceeds what most universities can pay faculty. This situation likely exists for other cross-disciplinary areas as well. It is expected that individuals with specialized knowledge in multiple domains are in high demand and have many lucrative and rewarding opportunities. Our solution at Cal Poly has been to leverage the expertise of interested and engaged faculty and to encourage "on-the-job training". Such an approach may not be appropriate for graduate programs, but an undergraduate program allows for faculty growth alongside the student growth.

ISSN:2149-8504 (online)

Acknowledgments

This work has been supported in part by the Cal Poly Office of Research and Economic Development and by the CAWIT (Center for Advancing Women in Technology) Technology Pathway Initiative (TPI) Award "Bioinformatics Cross-Disciplinary Studies Minor". The authors are grateful to a number of colleagues who have helped shape the Data Science, and Computation in Interactive Arts Cross-Disciplinary Studies Minors: Foaad Khosmood, Chris Lupo, Ignatios Vakalis, Zoe Wood, and Michael Haungs from the Department of Computer Science, and Hunter Glanz, Dennis Sun and Allan Rossman from the Department of Statistics. Additionally, we are grateful to the chairs of our respective departments' Curriculum Committees, Aaron Keen (CS) and Matt Carlton (Statistics).

References

Augsburg, T., & Henry, S. (2009). *The politics of interdisciplinary studies: Essays on transformations in American undergraduate programs*. McFarland &Jefferson, N.C..

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.

Dymond, R., Boyle, K., Beliveau, Y. J., Goss, R. C., Kumar, R., & Wiseman, E. (2015). Development of an interdisciplinary undergraduate program in real estate: Breaking Down University Silos. *Journal of Real Estate Practice and Education*, *18*(2), 141-162.

Finzer, W. (2013). The data science education dilemma. *Technology Innovations in Statistics Education*, *7*(2), 1-9.

Kim, J. Y., & Lee, C. K. (2016). An empirical analysis of requirements for data scientists using online job postings. *International Journal of Software Engineering and Its Applications*, *10*(4), 161-172.

Klein, J. T. (2009). *Creating interdisciplinary campus cultures: A model for strength and sustainability*. John Wiley & Sons.

Knight, D. B., Lattuca, L. R., Kimball, E. W., & Reason, R. D. (2013). Understanding interdisciplinarity: Curricular and organizational features of undergraduate interdisciplinary programs. *Innovative Higher Education*, *38*(2), 143-158.

Lafferty-Whyte, L. (n.d.). The rise of the Biologists: The changing face of the bioinformatics industry. Retrieved December 19, 2017, from http://stateofinnovation.com/the-rise-of-the-biologists-the-changing-face-of-the-bioinformatics-industry

Levine, A. G. (2014). An explosion of bioinformatics careers. *Science*, *344*(6189), 1303-1306.

Luke, D. A., Carothers, B. J., Dhand, A., Bell, R. A., Moreland-Russell, S., Sarli, C. C., & Evanoff, B. A. (2015). Breaking down silos: mapping growth of cross-disciplinary collaboration in a translational science initiative. *Clinical and Translational Science*, *8*(2), 143-149.

Parthasarathy, M., Forlani, D., & Meyers, A. (2015). The University of Colorado Certificate Program in Bioinnovation and Entrepreneurship: An Update and Current Status. *Journal of Commercial Biotechnology*, *21*(2).

Pournaras, E. (2017). Cross-disciplinary higher education of data science–beyond the computer science student. *Data Science*, (Preprint), 1-17.

Ranganathan, S. (2005). Bioinformatics education—perspectives and challenges. *PLoS Computational Biology*, *1*(6), e52.

Razzaq, J., Townsend, T., & Pisapia, J. (2013). Towards an understanding of interdisciplinarity: The Case of a British University. *Issues in Interdisciplinary Studies*, *31*, 149-173.

Tate, E. (2017, March 15). *Data Analytics Programs Take Off*. Retrieved December 19, 2017, from https://www.insidehighered.com/digital-learning/article/2017/03/15/data-analytics-programs-taking-colleges

ISSN:2149-8504 (online)

Van der Aalst, W. M. (2014). Data scientist: The engineer of the future. *In Enterprise Interoperability VI* (pp. 13-26). Springer, Cham.

Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education*, *15*(2), 180-187.

ISSN:2149-8504 (online)