BOARD # 294: Reinforcing curricular interventions in data science through an experiential learning internship program (NSF IUSE [GRANT NUMBER])

Prof. Petra Bonfert-Taylor, Dartmouth College

Petra Bonfert-Taylor received her Ph.D. in Mathematics from Technical University of Berlin (Germany), was a postdoctoral fellow at the University of Michigan and a faculty member in the Mathematics Department at Wesleyan University before accepting her current faculty position at the Thayer School of Engineering at Dartmouth. She was elected a Fellow of the Association for Women in Mathematics in the Class of 2020. Her mathematics research is in geometric function theory and discrete groups; she also has a strong interest in broadening access to high-quality higher education and pedagogical innovations that aid in providing equal opportunities to students from all backgrounds. This passion led her to design and create a seven-MOOC Professional Certificate on C-programming for edX for which her team won the ”2019 edX Prize for Exceptional Contributions in Online Teaching and Learning”. Previously she designed a MOOC ”Analysis of a Complex Kind” on Coursera. Petra is the recipient of the New Hampshire High Tech Council 2018 Tech Teacher of the Year Award, the Binswanger Prize for Excellence in Teaching at Wesleyan University and the Excellence in Teaching Award at Thayer. She recently co-designed and piloted a Foreign Studies Program focussed on green and sustainable engineering in collaboration with the German department at Dartmouth. At Thayer she furthermore leads an AAU funded Teaching Evaluation Project to develop, implement, and document a more effective and holistic teaching evaluation system.

Petra has served as Associate Dean for Diversity and Inclusion at Thayer since 2020. In this role she plans, leads and oversees diversity and inclusion efforts at Thayer and in coordination with other organizations internal and external to Dartmouth. Thayer was recently recognized with the ASEE Silver Diversity Award for our progress in increasing diversity and inclusion of our program.

Dr. Laura E. Ray, Dartmouth College

Dr. Ray is a professor of engineering sciences at the Thayer School of Engineering, Dartmouth College. She received her B.E. and Ph.D. degrees from Princeton Univ. and her M.S. degree from Stanford University. She is a co-founder of two companies. Her re

Scott Pauls, Dartmouth College

Reinforcing curricular interventions in data science through experiential learning (NSF IUSE 1917002)

Introduction

Expertise in data science is becoming a necessary skill for all STEM fields, with requirements ranging from data analysis and visualization to data collection, cleaning, management, and preservation. Proficiency in one or more of these areas in conjunction with domain knowledge within a core STEM discipline is rapidly becoming a key need for education and workforce development. To meet the need for STEM professionals with proficiency in data science, the NSF-sponsored DIFUSE project at Dartmouth has focused on integrating data science into STEM disciplines to enhance undergraduate student learning and preparation for the STEM workforce. The interdisciplinary approach, described in [1], develops data science modules for use in the classroom in introductory STEM and social science courses ranging from psychology and environmental studies to astronomy and engineering; to date, we have developed and disseminated over 20 such modules, both for Dartmouth courses and for courses at colleges and universities across the United States. The modules are available for download and use in a permanent repository [2]. We have analyzed the impact of the modules on student data science interest, beliefs, career aspirations, and self-efficacy [3] using a validated survey instrument [4]. We also assessed the impact of two workshops on the module development process on thirty faculty participants from across the country [5], finding growth in their skills, confidence and self-efficacy.

In addition to these early data science modules, a crucial element of the DIFUSE project is pairing students with practice in data science skills through experiential learning opportunities. To meet this need, we have created and deployed three unique experiential learning opportunities for audiences ranging from early STEM students to more advanced undergraduates. First, we create teams of undergraduates and graduate students to support data science module development for early STEM courses within the context of a term-long team project, beginning with definition of learning objectives in conjunction with the course instructor, and culminating with a finished, ready-to-deploy module and associated learning materials. Second, we have supported 31 students to-date through experiential learning internships incorporating data science, either within a research setting or with a nonprofit organization. These opportunities benefit early students who may have difficulty obtaining such internships in for-profit companies due to lack of experience. Finally, we support students in Dartmouth's Digital Applied Learning and Innovation (DALI) Lab in team-based development of mobile applications, websites, virtual & augmented reality, and digital installations. The experiential learning opportunities serve both to cement classroom learning and to allow students to explore careers in data science. This paper focuses on the experiential learning models and outcomes.

Experiential Learning Through DIFUSE Module Development

A central component of the DIFUSE project is the development and implementation of data science modules, which provide hands-on opportunities to apply theoretical concepts in real-world contexts. Each module is developed with the help of a team of undergraduate and

graduate student interns. Working on such module development teams not only enhances interns' technical expertise but also fosters interpersonal growth and helps them clarify their career aspirations. To evaluate the impact of module development on interns, we use Likert-scale surveys (1 = low, 5 = high) to measure growth in confidence, interest, and skill development across three key domains: professional and interpersonal skills, technical expertise, and career alignment. Additionally, qualitative case studies provide rich insights into the individual experiences and outcomes of DIFUSE interns highlighting the diverse impacts of the program on participants. Together, these data highlight the program's significant role in preparing students for the demands of a data-driven, interdisciplinary workforce.

Interns' responses to surveys reveal three core areas of growth that illustrate DIFUSE's transformative impact: development of professional and interpersonal skills, strengthening of technical knowledge, and achievement of personal and professional goals. These interconnected outcomes demonstrate how the program holistically supports interns' professional development.

Development of Professional and Interpersonal Skills: Collaborative projects help interns build essential workplace skills such as teamwork, leadership, and communication. Student A reflects, "The program improved my teamwork skills and gave me useful insight into working with clients to achieve a working product." Student B highlights her growth in leadership, stating, "I gained so much confidence in my leadership, initiative, and collaboration abilities." Student C adds, "I gained confidence in presenting curriculum to multidisciplinary leadership teams." These qualitative insights are supported by survey data, which show that confidence in professional collaboration increased from a mean of 3.2 before participation in the program to 4.3 afterward. By fostering these interpersonal skills, DIFUSE prepares interns to thrive in interdisciplinary, team-based environments.

Strengthening Technical Knowledge Through Applied Learning: Hands-on opportunities allow interns to deepen their technical expertise. Projects involving Python programming, data visualization, and curriculum development allow interns to apply their knowledge in practical, transferable ways. Student C remarks, "I learned Python for the first time, which was an exciting addition to my skill set," while Student B notes, "I gained exposure to many different aspects of data science, which broadened my understanding of its applications." Student D, who adapted modules for healthcare contexts, shares, "I reinforced my personal interest in data science and learned to use data tools effectively." Survey results further illustrate this growth, with confidence in technical skills such as data visualization rising from a mean of 3.5 to 4.2. These experiences highlight DIFUSE's emphasis on applied learning, enabling interns to tackle real-world challenges with confidence.

Achievement of Personal and Professional Goals: Beyond technical and interpersonal growth, the program plays a pivotal role in helping interns align their career aspirations with academic and professional opportunities. Student A states, "The program clarified my career goals and made me more comfortable working on data science projects." Similarly, Student E reflects, "The program reinforced my personal interest to pursue data science education formally at some point in my career." Student F emphasizes the program's influence on his academic trajectory, noting, "DIFUSE encouraged me to take more data science classes at Dartmouth." Survey data shows that career clarity scores increased from a mean of 3.4 to 4.4 through the program,

illustrating how DIFUSE supports interns in identifying and pursuing meaningful professional pathways.

DIFUSE offers a model for experiential learning in STEM education. By focusing on module development, the program bridges the gap between academic preparation and professional application, equipping students with the skills, confidence, and adaptability needed to address the multifaceted challenges of modern industries. Institutions adopting similar frameworks can foster innovation, accessibility, and workforce readiness, ensuring that graduates are prepared to thrive in an increasingly data-driven world.

Experiential Learning Through Internships With Research and Non-Profit Organizations

To extend the impact of the DIFUSE project we create another path for students to develop data science skills through a collaboration with our institution's Center for Professional Development (CPD). Using CPD's communication and application platform, we solicited applications for internship funding in support of internships that were data intensive within non-profit organizations. To date, we've supported 31 students, and we've focused our support on projects that had explicit support structures for students that extended their data science, communication, and teamwork skills.

Internships cover an extremely diverse set of opportunities for students. Projects touch different aspects of data science - from data collection and cleaning all the way through experiment design, analysis, interpretation, and communication. Students engaged in projects in numerous fields, including health care, neuroscience, ecology, agriculture, public policy, criminal justice, plasma physics, orthopedics, legal informatics, opioid use disorder, and more.

For example, one student interned at the National Spine Health Center tackling the difficult problem of teasing out trends from the unstructured data within medical records. He traced demographic trends and impacts of treatment paths to build models to improve patient outcomes. Another student worked on a neuromechanics project at a lab at the University of Colorado, Boulder where she applied machine learning techniques to motion capture data to help uncover aspects of the brain's processes for controlling motion. A third student is working in MIT's Media lab on a project using transformer artificial intelligence techniques on musical data to create collaborative human and AI musical compositions. A fourth student worked on the Experimental Agroforestry Transition (EAT) plot at Dartmouth to study how structural and functional complexity in agroforestry designs influence plant-soil feedback and successional dynamics. She identified and collected data on weed and crop species abundance and composition through the growing season in relation to experimental treatments, with a focus on functional trait effects. Additional soil and species data were collected on partnering New England farms that have implemented their own agroforestry system designs.

Upon completion of their internships, students report satisfaction with their experience as well as stronger data science skills, albeit with differing degrees of increase.

Data Science component to DALI Lab Extracurricular Learning

DIFUSE also provides a transformative platform for students to engage in experiential learning through the Digital Applied Learning and Innovation (DALI) Lab. DALI is an extracurricular program at Dartmouth. At DALI, students work on interdisciplinary teams to tackle real-world problems, combining design, technology, and user experience to translate research into impactful digital tools. By participating in projects funded by DIFUSE, such as the Pine Beetle Project and the Smart Microscope Project, students not only apply their academic knowledge to practical challenges but also gain hands-on experience in data science and tool development. These projects help students see how their work directly affects communities, industries, and the environment, offering invaluable insights into both the technical and human-centered aspects of innovation.

The ***Smart Microscope Project*** (with Dr. Aravindhan Sriharan, Dermatopathologist and Professor of Pathology and Laboratory Medicine, Dartmouth) exemplifies how students learned to bridge the gap between traditional workflows and modern technological advancements. While working on this project, students explored the fundamentals of medical diagnostics and the limitations of traditional microscopes on telepathology. They then applied data science to design a tool that integrates machine learning (ML) without disrupting the established practices of dermatopathologists. Through this process, students gained expertise in creating state-of-the-art user interfaces and integrating AI algorithms in real-time systems. More importantly, they saw the direct impact of their work: a tool that enhances diagnostic accuracy, reduces costs, and saves time for patients. This hands-on experience not only deepened their understanding of AI in medicine but also taught them how to develop user-friendly solutions for high-stakes fields like healthcare. The work is published in [6].

The ***Pine Beetle Project*** (with Dartmouth Matthew Ayers, Professor of Biological Science, Co-chair, Graduate Program in Ecology, Evolution, Environment and Society at Dartmouth) offered students a chance to engage with an urgent environmental issue while applying advanced predictive modeling techniques. Students learned about the biology and ecological impact of the southern pine beetle, gaining insights into how climate change exacerbates pest outbreaks. They worked with Prof. Ayres to develop a statistical model to predict outbreaks and designed a web-based interface for forest managers. This experience exposed students to the complexities of large-scale data collection, modeling, and visualization. By working on a tool that could streamline data input, enhance predictive accuracy, and serve as an educational resource, students saw how their technical skills could contribute to sustainable forest management and public awareness. These projects exemplify how the DIFUSE-DALI partnership not only equips students with advanced data science skills but also instills a sense of purpose and confidence in creating tools with meaningful, real-world impact.

Conclusion

The DIFUSE project at Dartmouth is a pioneering initiative in equipping STEM students with essential data science skills through an innovative blend of classroom instruction and experiential learning. As evidenced by the program's multifaceted approach—involving module development, internships, and partnerships—students are not only cultivating technical expertise but also building critical interpersonal and professional skills. DIFUSE's impact is apparent in

the qualitative feedback and quantitative survey results gathered from participants, which underscore notable increases in confidence across three dimensions: technical knowledge, professional collaboration, and career clarity. Students have embraced opportunities to apply their academic knowledge to practical challenges, ranging from medical diagnostics to ecological modeling. These experiences ensure that graduates are better prepared to navigate interdisciplinary, data-driven sectors.

As the demand for data-literate professionals continues to grow, initiatives like DIFUSE will be instrumental in closing the skills gap. We encourage other educational institutions to cultivate similar partnerships and learning models, thereby advancing workforce readiness, fostering innovation, and ensuring that the next generation of professionals is equipped to thrive in an increasingly data-centric world.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Number DUE-1917002. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

References

1. Bonfert-Taylor, P., Ray, L., Pauls, S., Loeb, L., Sankey, L., Busch, J., & Hickey, T. (2022, August). Infusing Data Science into the Undergraduate STEM Curriculum. In *2022 ASEE Annual Conference & Exposition*.
2. Dartmouth DIFUSE: The github home for Dartmouth College's Data Science Infused in STEM Education Project. https://github.com/difuse-dartmouth
3. Kursav, M. N., Pauls, S. D., Ray, L., Bonfert-Taylor, P., & Loeb, L., (Under-review) Intersections of identity and interventions in the impact on student attitudes towards data science. J*ournal of Statistics and Data Science Educatio*n.
4. Li, Z., Pauls, S., Szendey, O. (2018) Survey of Attitudes towards Data Science, doi: 10.5281/zenodo.6884050
5. Kursav, M. N., Pauls, S. D**.**, Bonfert-Taylor, R., Loeb, L., Ray, L. (2025) Accelerating Change of Practice through Targeted Professional Development in Data Science. *Journal of Statistics and Data Science Education*, 1–25. 10.1080/26939169.2025.2479012
6. Aller, Raymond D. and Winstein, D. Smart microscope evokes look and feel of the traditional instrument, CAP Today Newsbytes, https://www.captodaymag.com/captoday/august_2024/MobilePagedArticle.action?articleId=2002566#articleId2002566.