Purdue University Libraries and School of Information Studies

DATA SCIENCE

1,00.00



Libraries and School of Information Studies

Beth McNeil DEAN OF LIBRARIES AND SCHOOL OF INFORMATION STUDIES



Here at Purdue University Libraries and School of Information Studies, we are proud of our longstanding commitment to collaboration with Purdue faculty and staff through a wide variety of data-related initiatives.

As we enter a new decade, I am pleased to showcase some of the innovative and forward-thinking data science projects spearheaded by our faculty and staff. These initiatives directly support Purdue's mission of graduating students with the ability to understand, evaluate, and integrate information – including data – to develop the invaluable problem-solving, critical thinking, and decision-making skills necessary to achieve a lifetime of professional and personal success.

As you explore some of our collaborative projects, you will quickly discover the broad impact that Purdue University Libraries and School of Information Studies has on our growing enrollment of 40,000+ students. Our efforts are interdisciplinary, creative, and grounded in real-world applications. We

mentor undergraduate students and teach them how to effectively communicate their research findings. We provide training and support for faculty seeking to integrate data science into their courses.

The geospatial training that we develop will help meet an anticipated expertise shortage in the United States, and our new programs in Critical Data Studies, Digital Humanities, and the Engineering in the World of Data Learning Community provide students with opportunities to engage more deeply with data-driven projects, developing proficiencies that will make them smart data consumers and highly marketable job candidates.

This is an exciting time as we continue to evolve with technological advances, the 21st century educational landscape, and the growing demands of industry knowledge.

As always, thank you for your support. We couldn't do it without you!

Dean of Purdue University Libraries and School of Information Studies Esther Ellis Norton Professor of Library Science



Image Credits: financial graph (Cover, p.1-2) ©monsiti, Getty Images; Vector people evolution from monkey to robot (p.3) ©Sabelskaya, Getty Images; Smartphone (p.3) ©Smashicons; Mouse (p.3) ©srip; Shopping Cart, Click (p.3) ©Freepik; Signal (p.4) ©Freepik; Analysis, Monitor, Git (p.6) ©Freepik, flaticon.com; Network connection and global data exchanges on blue planet earth- elements of this image furnished by NASA (p.7-10) ©ipopba, Getty Images: Satellite (p.9) ©Freepik, flaticon.com; Rubik (p.15-17) ©Freepik, flaticon.com; circuit board in the Cyborg brain, Artificial intelligence of digital human. (p.15) ©hakule, Getty Images; Andrew J. Feustel (p.21) NASA.gov: Data analysis line icons in round shape isolated vector illustration (p.25) ©FANDSrabutan, Getty Images; Businessman analyzes page data (p.25) ©z_wei, Getty Images; DNA test infographic, Genome sequence map, chromosome architecture and genetic sequencing chart abstract data vector illustration (p.26-28) ©Tetiana Lazunova, Getty Images; Hand drawn physics formulas Science knowledge education (p.34, 36) ©undrey, Getty Images: Flat Business People Climbing Up The Stairs Career Ladder with Characters. Team Work, Partnership, Leadership Concept, Vector illustration (p.35) ©invincible_ bulldog, Getty Images; Artificial intelligence (p.37-42) Opiranka, Getty Images

Purdue University Libraries and School of Information Studies **DATA SCIENCE**

PAGE 3-6

ILS COURSE AIMS TO CULTIVATE CRITICAL DATA CONSUMERS AND RESEARCHERS

PAGE 7-10

A HUB FOR GEOSPATIAL INFORMATION STUDIES AT PURDUE

PAGE 11-14 IMPACT DATA SCIENCE

PAGE 15-18 CRITICAL DATA STUDIES AT PURDUE

PAGE 19-22 LIVING, LEARNING, AND DOING DATA SCIENCE

PAGE 25-28 DATA SCIENCE PROJECT CRITICAL FOR PURDUE EDUCATION

PAGE 29-30 OUR COURSES TEACH EFFECTIVE DATA-SUPPORTED COMMUNICATION

PAGE 31-36 A CURE FOR SCALING UNDERGRADUATE RESEARCH

PAGE 37-42 ADVANCING DIGITAL HUMANITIES AT PURDUE

Publication Staff:

Teresa Koltzenburg, Writer/Editor Ayn T Reineke, Writer/Editor Jarrod Hurt, Designer/Creative Services Manager

Images from Getty Images, NASA, Purdue University Marketing and Media, Purdue University Libraries and School of Information Studies, Teresa Koltzenburg, John Underwood



ILS COURSE AIMS TO CULTIVATE CRITICAL DATA CONSUMERS AND RESEARCHERS



by Megan Sapp Nelson, Professor

aculty at Purdue University would be missing the mark by a long shot if they were not incorporating big data concepts into their courses to demonstrate how critical it is today.

"Big data is transforming our society," noted Megan Sapp Nelson, Professor and Science and Engineering Librarian in Purdue University Libraries and School of Information Studies. "With every click on the Internet and every transaction we create, even as we move through geographic space with our cell phones, we are creating new data about ourselves. Everything we are doing in time today is quantified in some way as data."

So it is not surprising that our innovative Purdue faculty across a variety of disciplines are incorporating instruction about big data and its four dimensions volume, variety, velocity, and veracity—into their courses. The Integrative Data Science Initiative (IDSI), too (more information available on p. 12), is advancing both Purdue students' and faculty members' understanding of how data impacts every discipline, every field of study—and, as Sapp Nelson stated, our society in general.

"We are teaching Purdue students to be critical consumers and researchers of data in their relevant disciplines. So when they graduate, they will be able to apply this knowledge to both their professional and personal lives," she added.

Sapp Nelson's graduate course which she co-teaches with Assistant Professors Pete Pascuzzi and Chao Cai— ILS 595: Data Management at the Bench, began as an initiative among faculty in the Purdue College of Agriculture and collaborators a few years ago.

ILS STANDS FOR Information and Library Science



VARIETY

VELOCITY

VERACITY

ILS COURSE AIMS TO CULTIVATE CRITICAL DATA CONSUMERS AND RESEARCHERS CONTINUED FROM PAGE 4

"Our initial goal was to prepare graduate students to be ready to manage the increasing volume of data produced in agriculture," she explained. "We then expanded the course to serve all students working in disciplines primarily producing numerical or quantitative data. We christened and re-conceptualized the course as, 'Data Management at the Bench.' The material in the course provides a comprehensive, systematic introduction to data management, within the context of managing data that is of most interest to them."

Graduate students and postdoctoral student researchers creating numerical or quantitative data in their individual disciplines are eligible to enroll in ILS 595. According to Sapp Nelson, this has led to growing enrollment from diverse disciplines across campus, with the majority of students studying in the College of Science, College of Agriculture, and Purdue Polytechnic Institute.

EVERYTHING WE ARE DOING IN TIME TODAY

"In the course, we start with the basic principles, such as file naming and file organization, then move on to more advanced skills, such as data wrangling, cleaning, and visualization in a high-performance computing environment," Sapp Nelson noted. "This class is also flexible in learning objectives, so if students have specific skills they want to learn—such as mapping geo-coordinates in the program R—the instructor team is willing to adjust the syllabus to fit instruction to ensure that students get information on specific content."

Because of the flexible design of the course, Sapp Nelson said students' learning outcomes can encompass a wide array.

"Outcomes include students knowing how to: efficiently use a computer as their primary data management tool based on



knowledge of how files and data are stored on Windows, Mac, and Unix computer systems; devise efficient strategies using tools—such as Excel, R, and Git—to collect, store, organize, visualize, analyze, and publish their research data; discover a wide range of public data repositories so they can identify, acquire, explore, clean, and incorporate published data into their own research projects; prepare datasets for publication by identifying appropriate repositories, choosing the best format for data sharing and assigning sufficient metadata; and devise and implement a data-management plan that complies with one or more major government funding agencies," she explained.

ILS 595 is just one of many courses offered through Purdue University Libraries and School of Information Studies.

IS QUANTIFIED IN SOME WAY AS DATA.

A HUB FOR GEOSPATIAL INFORMATION STUDIES AT PURDUE

Since the beginning of this century (and even before), geospatial information has played an important role in agriculture and engineering, two fields that have easy-to-discern reasons why using this type of data is key.

But, in our increasingly datadriven world, the use of geospatial information systems (GIS) has significant implications for business, humanities, the social sciences, health care, and a bevy of other fields and disciplines. As the use of geospatial information has grown, so has the need for professionals who have geospatial information skills. In fact, U.S. Department of Labor officials predict that geospatial technologies will be one of the major areas of job growth in the coming decade.

This need is why Purdue University Libraries and School of Information Studies Associate Professor and GIS Specialist Nicole Kong and a team of faculty from seven other colleges at Purdue have come together to identify gaps in geospatial education here. Through the project, "Integrating Geospatial Information Across Disciplines," Kong, who serves as the principal investigator (PI), said their goal is to develop a proposal to create and implement a graduate certificate program so students can earn a geospatial information certification.

"As of today, Purdue does not have a harmonized effort to offer a geospatial data training program, although related courses exist within the departments where instructors are available," she explained. "The team behind our proposal is comprised of faculty who teach these courses, as well as faculty from other departments who have expressed a need for comprehensive geospatial training at Purdue." According to Kong and her research team, systematical training in geospatial information skills not only allows students to integrate data from various domains, but it also provides them with additional analytical methods to reveal relationships that may not be obvious in traditional analyses. Purdue graduates with GIS knowledge could easily have a competive advantage in the job market.

> Kong also noted that "the interdisciplinary nature of Purdue University Libraries and School of Information Studies provides a great opportunity for the Libraries to become a centralized place to coordinate campus-wide training in geospatial information studies."

A HUB FOR GEOSPATIAL INFORMATION STUDIES AT PURDU CONTINUED FROM PAGE 8



In their proposal for "Integrating Geospatial Information Across Disciplines," the team pointed to an Urban and Regional Information Systems Association (URISA) survey result that shows the average salary for an individual with related certification is \$10,000 more per year.

IDSI Support

In Fall 2019, through the support of Purdue's Integrative Data Science Initiative (IDSI), Kong and her team were given the green light and support to roll forward with the task of developing a proposal for the graduate certificate program.

"Over the next year of this project, we will interview faculty members and host workshops to coordinate the current teaching efforts and find the gaps in geospatial education on campus," she said. "Essentially, we are proposing to further develop geospatial information education curriculum at Purdue to help students harness the skills to analyze and manage the ever-increasing amount of geospatial data. We will build the faculty and support team to develop a geospatial information graduate certificate program by developing advanced courses and seminars, and reorganizing existing courses."



Aishwarya Chandrasekaran (Chennai, India), a master's degree candidate in forestry and natural resources, presents her research project, "UAS Imagery for Automatic Measurement of Tree Canopies in a Red Oak Plantation" at the Purdue GIS Day Conference Nov. 8, 2019.

Kong also noted that "the interdisciplinary nature of Purdue University Libraries and School of Information Studies provides a great opportunity for the Libraries to become a centralized place to coordinate campus-wide training in geospatial information studies."

Purdue graduates with GIS knowledge could easily have a competitive advantage in the job market.

"In addition, we would also like to recognize students who have already made efforts in this area. Purdue has a large group of graduate students from multiple disciplines who have received outstanding training in analyzing geospatial information, but who are not formally recognized with a certificate or minor right now. For example, in the ArcGIS Online (a GIS platform), new users from Purdue increased by more than 1,000 from 2015-18. A coordinated effort and a certificate program can benefit these students and provide them with formal recognition of their accomplishments."

Other investigators on the project include: Forestry and Natural Resources Professor Bryan Pijanowski, Associate Professor Songlin Fei, and Assistant Professor

Brady Hardiman (College of Agriculture); Civil Engineering Professor Jie Shan (College of Engineering); Agricultural and Biological Engineering Associate Professor Dharmendra Saraswat (College of Agriculture); Anthropology Associate Professor Ian Lindsay (College of Liberal Arts); Professor Michael Fosmire (Libraries and School of Information Studies); Pharmacy Practice Assistant Professor Ephrem Abebe (College of Pharmacy); Department of Computer Graphics Technology Assistant Professor Vetria Byrd (Polytechnic Institute); Associate Professor and Director of the Data Science Consulting Service Guang Lin; Director of Research Services and Support, IT Research Computing Preston Smith; and Center for Instructional Excellence Instructional Developer Erica Lott.

DATA SCIENCE

When the Integrative Data Science Initiative (IDSI) was announced, Purdue University Libraries and School of Information Studies Associate Professor Clarence Maybee saw the initiative as an opportunity to create a sustainable and scalable course development process for integrating data science into undergraduate courses.



In Fall 2019, this process was supported when Maybee's team project proposal, "IMPACT Data Science Education: Preparing Undergraduates to Lead into the Future," was funded through the IDSI's second round of research funding.

"Knowing that Purdue is interested in graduating undergraduate students with data science skills that will enable them to lead into the future, we brought together the programs on campus that are already doing this type of work," Maybee explained. "Many



Libraries and School of Information Studies

PREPARING UNDERGRADUATES TO LEAD INTO THE FUTURE



(Photo Courtesy of Purdue University Marketing and Media/John Underwood)

of us leading the Instruction Matters: Purdue Academic Course Transformation program, commonly known as IMPACT, have been working with Purdue instructors since 2011 to make courses more engaging and student-centered. The Data Science Consulting Service works with instructors specifically to help them integrate data science into Purdue courses. IMPACT and the Data Science Consulting Service are the right partners, as we all have experience working to develop curricula."

With the goal to develop a model for creating coursework that integrates data science into undergraduate courses, Maybee, the principal investigator (PI), and his team have recruited eight Purdue University faculty members

IMPACT DATA SCIENCE CONTINUED FROM PAGE 12



(Photo Courtesy of Purdue University Libraries and School of Information Studies)

who are tasked with designing innovative and engaging data science coursework. New faculty members involved are required to participate in the IMPACT program and related activities.

"Our research team will collect data from the participants throughout the project to help determine what supports are the most beneficial to instructors to develop data science coursework." The six people on the team have a range of expertise that collectively supports the success of the IMPACT Data Science Education project. All team members have pedagogic knowledge," Maybee said. "Our evaluation of the project will be used to inform next steps towards developing an ongoing program to meet the goal."

As the PI, Maybee's role is to bring everyone together to set objectives, determine activities, and measure the success of the project. He said everyone on the team is leveraging their expertise to create project activities, which include four data science education working meetings in early 2020 and a Data Science Symposium in late spring. In addition to Maybee—who has studied how students use information and data to learn in various disciplinary contexts and has been a leader in the IMPACT program since 2012—the team includes:

- Jason Fitzsimmons, Educational Developer, Center for Instructional Excellence
- Guang Lin, Associate Professor and data scientist, Mathematics, Statistics, and School of Mechanical Engineering
- Yixuan Sun, graduate assistant and data scientist
- Wei Zakharov, Assistant Professor and Chao Cai, Assistant Professor, both Purdue University Libraries and School of Information Studies faculty members involved in IMPACT, who also teach courses in data management and the use of data science tools

LEARN MORE ABOUT

"IMPACT Data Science Education: Preparing Undergraduates to Lead into the Future" at https://sites.lib.purdue.edu/dse



Examining Democratic, Participatory Approaches to Al and Data Management

ritical Data Studies (CDS) is an interdisciplinary field that addresses the ethical, legal, socio-cultural, epistemological, and political aspects of data science, big data, and digital infrastructure. Purdue University Libraries and School of Information Studies faculty play a lead role in the development of CDS 0 initiatives at Purdue University, including the development of the monthly Critical Data Studies Seminar and Lecture Series, developing undergraduate and graduate-level courses, launching the CDS-Data Mine Learning Community, and collaborating with the Center for Science of Information.

Critical Data Studies—Data Mine Learning Community

The CDS-Data Mine Learning Community is one of more than 20 different Data Mine-branded residential learning communities located in Hillenbrand Hall. The CDS-Data Mine Learning Community lives up to the motto of "Data Science for All" by placing particular emphasis on democratic and







Bethany McGowan Assistant Professor







RITICAL DATA STUDIES AT PURDUE CONTINUED FROM PAGE 16



participatory approaches to algorithm design and responsible data management, curation and dissemination. Students engage in debates surrounding "technopolitics," such as the desire for explainable artificial intelligence (AI), algorithm and data justice initiatives, and critiques of discriminatory design. We also explore various attempts to disrupt power inequities exacerbated or brought on by a shift to a data economy. Students have the opportunity to engage in handson, community-based participatory design research and to develop models for non-extractive and justice-centered approaches to data collection, sharing, and use

practices. This may take the form of data and design hack-a-thons, directed or independent research projects, and/or public policy work. Opportunities to engage the larger campus community through active involvement in the Critical Data Studies Seminar and Lecture Series forms a key component of this Learning Community.

The CDS-Data Mine Learning Community is led by Assistant Professors Kendall Roark and Bethany McGowan, and Associate Professor Laura Zanotti (Anthropology), in consultation with Professor Mark Ward (Statistics). In addition, the CDS-Data Mine Learning Community has benefited from curricular collaborations with Assistant Professor Danielle Walker, and Faithe Day (African American Studies), faculty from the Honors College, and members of the Critical Data Studies Collective at Purdue.

Critical Data Studies and the Center for Science of Information

Purdue University Libraries and School of Information Studies faculty are also collaborating with the Center for Science of Information (CSol). The CSol is in its 10th year of grant funding from the National



We explore various attempts to disrupt power inequities exacerbated or brought on by a shift to a data economy." -Kendall Roark

Science Foundation (NSF) Science and Technology Center with the goal to embed critical data studies frameworks within informal graduate data science education. These joint projects include the week-long "CSol Summer Workshop on Critical Data Visualization" in 2019, a collaboration between Assistant Professor Kendall Roark, Assistant Professor Vetria Byrd (Purdue Polytechnic Institute), and CSol Director of Education Brent Ladd. This workshop aims to introduce graduate students to team science, help them build data science skills, and help them

learn how to implement ethical and just data visualization practices. In addition, Roark has worked with anthropology doctoral candidate Madisson Whitman (pictured above) to develop a series of Critical Data Studies learning modules for the CSol Learning Hub, which will launch during the Spring 2020 semester.

LEARN MORE ABOUT

lecture series, courses, and the learning community, information can be found at purdue.edu/critical-data-studies and datamine.purdue.edu/critical-data

LIVING, LEARNING, AND DOING DATA SCIENCE

By Michael Witt Interim Associate Dean for Research Associate Professor



Purdue Men's Basketball Coach, Matt Painter (far right), and the instructors and students in the Engineering in the World of Data (EWD) Learning Community, Fall 2018, in Mackey Arena. (Photo courtesy of Teresa Walker, Purdue School of Engineering Education.)

This year, 97 first-year students in the Engineering in the World of Data (EWD) Learning Community live together in Harrison Hall and share a passion for data science and engineering. In the EWD Learning Community, which I, along with Sean Brophy (Associate Professor, Purdue School of Engineering Education) lead students that enroll in five data science-themed courses together as a cohort and engage in fun, exciting extracurricular activities. In Fall 2018, students watched the movie "Moneyball" and learned about the application of data science to sports. The following day, Purdue Men's Basketball Coach, Matt Painter, served as a guest lecturer and shared his experience with analyzing data for recruiting, scouting opponents, and real-time decision-making during games. Based on this and other activities, our EWD Learning Community teaching team received the 2019 Academic Connection Award from Residential Academic Initiatives for teaching and engagement in Spring 2019.

ENGINEER

IN THE WORLD OF DAT LEARNING COMMUNIT



LIVING, LEARNING, AND DOING DATA SCIENCE CONTINUED FROM PAGE 20



More recently, in Fall 2019, the students in the EWD Learning Community worked on a simulated problem to identify and provide the data that NASA would need to fix the Hubble Space Telescope after it collided with debris in space. Students presented their solutions and enjoyed lunch with Purdue alumnus and NASA astronaut Drew Feustel.

Assistant Professor Nastasha Johnson, who co-teaches ILS 295: Developing Your Data Mind, noted that she appreciates the opportunity to integrate topics across different courses and bring data science to life through activities outside of the classroom.

"The year-long EWD Learning Community provides students with the opportunity to build friendships while developing a deeper understanding of data," Johnson said. "After the year in our community, they leave with both."



Purdue alumnus and NASA astronaut Drew Feustel (BS'89, MS'91, HDR'18, end of table, left) talks with EWD Learning Community students at the Astronaut Reunion, held during Purdue's 2019 Homecoming celebration in October.



Drew and Indira Feustel (MS'91, front row, center) pose with instructors and students in the 2019-2020 EWD Learning Community.

Thomas S. and Harvey D. Wilmeth Active Learning Center



Libraries and School of Information Studies

(Photo Courtesy of Purdue University Marketing & Media/John Underwood)

Data Science Project Critical for Purdue Education

Through his teaching and consulting in data science, Purdue University Libraries and School of Information Studies Assistant Professor Pete Pascuzzi has learned a lot about the distinctive range of Purdue University



(Photo Courtesy of Purdue University Marketing and Media)

students' and faculty members' needs when it comes to data science and data management.

"I teach courses in bioinformatics and data management, but many of my consultations with learners (students and/ or researchers and/or faculty) are independent of the classes I teach," Pascuzzi explained. "From these consultations, it became clear to me that learners have diverse needs in data science, including guestions about data acquisition, cleaning, analysis, visualization, documentation, or computing. Many credit classes at Purdue cover these topics, but it is not always possible for learners to enroll in these classes because of time constraints."



Pascuzzi also noted the data science infrastructure at Purdue includes a variety of resources—including ITaP (Information Technology at Purdue) computer labs, Research Computing Clusters, the Purdue University Research Repository (PURR), and software development platform GitHub. Through the project "Building a Data Science Education Ecosystem Resource Collection," or DSEERC, Pascuzzi and his colleagues are creating a collection that includes curated instructional material produced by and for data science efforts on campus. Pascuzzi is the principal investigator (PI) of the project, which was recently

supported by the Integrative Data Science Initiative (IDSI) at Purdue.

According to Pascuzzi, he and his team are developing and collecting material to share while developing the infrastructure to facilitate this sharing.

They will start sharing material this spring with the Data Mine Learning Community students and specific courses, and they anticipate a broader roll out for Fall 2020. "Material in the collection is comprised of interactive tutorials, videos, lessons and datasets, with an emphasis on modular content that can be used in existing courses or as needed by students.

DATA SCIENCE PROJECT CRITICAL FOR PURDUE EDUCATION CONTINUED FROM PAGE 26

Learners at Purdue will be able to search for material, and instructors will be able to search for content they can use or adapt for their classes," he said. "The collection will include faceted search capabilities to allow learners to quickly identify appropriate resources. In addition, downloads will be tracked, and the repository will include a feedback mechanism, so contributors to the collection can track their impact and improve their materials."

"The collection will include faceted search capabilities to allow learners to quickly identify appropriate resources. In addition, downloads will be tracked, and the repository will include a feedback mechanism, so contributors to the collection can track their impact and improve their materials."

- Pete Pascuzzi



(Photo Courtesy of Purdue University Libraries and School of Information Studies)

Pascuzzi said the project leverages the expertise of faculty in Purdue University Libraries and School of Information Studies, staff in Research Computing, and the students and instructors in the Data Mine Learning Community. Pascuzzi's co-Pls on the project include Michael Witt (Interim Associate Dean for Research and Associate Professor), who is coordinating efforts with PURR; Preston Smith (Director of Research Services and Support, Research Computing), who is coordinating efforts with Research Computing, including infrastructure upgrades; and Mark Ward (Professor of Statistics and Associate Director of Actuarial Science), who is coordinating efforts with the Data Mine Learning Community, which introduces students to concepts in data science. As PI, Pascuzzi will coordinate efforts among those contributing to the DSEERC.

OUR Courses Teach Effective Data-Supported Communication

By JJ Sadler, Associate Director

The Purdue Office of Undergraduate Research (OUR) developed and offers a first-of-its-kind online curriculum for both prospective and current undergraduate researchers. This course series delivers a comprehensive and crossdisciplinary view of research and creative endeavors that no other university offers. One major theme across this four course curriculum is the value of effective datasupported communication to various audiences.

Two of these courses—intended for current researchers—focus on communicating a student's research project and data to general, educated audiences, as well as to field experts. While participating in these courses, students strengthen their abilities to convey concisely and effectively to these diverse



audiences how their data support a research project's significant findings. Student researchers learn to connect their conclusions with their project data, to create data visualization components to convey the meaning of their results, and to draw in the listeners for maximum engagement during brief interactions.

These courses address writing a research abstract, constructing an oral research pitch, and building on communication skills through the development and presentation of a research poster with effective (Photo Courtesy of OUR Outreaching Intern Elizabeth Coats)

use of static visual and textual representations of a project's background, methods, results, and conclusions. Both courses use iterations of these assignments for students to apply peer and instructor feedback to their final products. Ultimately, these project assignments create usable content, such as an abstract and research poster for conferences and research pitches for professional exchanges. It is inspiring to observe these student researchers contemplate the next steps and broader impacts of their research projects as they

learn more about their fields and the knowledge gaps they can pursue in the future.

The Office of Undergraduate Research's mission is to serve as a central resource to promote and expand experiential learning for undergraduate students through research experiences with skilled research mentors.

LEARN MORE ABOUT

the online course series and various events hosted by OUR can be found online at purdue.edu/undergrad-research

A CURE FOR SCALING UNDERGRADUATE RESEARCH

By Teresa Koltzenburg



Alan Wright, Rafael Lang and Emily Kincaid

The Office of Undergraduate Research's CURE-Purdue Program provides faculty with training, support, and resources to incorporate research into coursework to involve more students in research-helping them hone their data science skills, and much more, in their undergraduate years.



n Fall 2019, Purdue University students Emily Kincaid and Alan Wright, along with close to 60 other undergraduate students (mostly first years), were conducting research—through data analysis—to detect dark matter. Both physics majors and sophomore-level students, Kincaid (Zionsville, IN) and Wright (Schererville, IN) were working on this research through a Physics course taught by Associate Professor of Physics and Astronomy Rafael Lang, which is one of the course-based undergraduate research experiences (CURE) in the Office of Undergraduate Research's CURE-Purdue Program.

Lang's undergraduate Physics course focuses on his particular research interest, the search for dark matter particles using the massive XENON1T experiment, which is located under a mountain in central Italy. On Lang's faculty web page, he ponders: "What is the universe made of? All we know today is that we don't know what most of the stuff in the universe really is. We have names for it, calling it 'dark matter' and 'dark energy,' but we don't know its true nature."

Through the Office of Undergraduate Research's CURE-Purdue program, Lang is able to involve many undergraduate students in real research and data science analysis, who may, one day, help find the true nature of dark matter. I have been building undergraduate research since I started at Purdue eight years ago, and it's grown more and more. But with a dozen or so students each semester I hit a ceiling, and beyond that, scaling was difficult. The resources provided through the Office of Undergraduate Research really helped me grow beyond that, to the 50, 60 students we have now.

A CURE FOR SCALING UNDERGRADUATE RESEARCH CONTINUED FROM PAGE 32

"I have been building undergraduate research since I started at Purdue eight years ago, and it's grown more and more. But with a dozen or so students each semester I hit a ceiling, and beyond that, scaling was difficult. The resources provided through the Office of Undergraduate Research really helped me grow beyond that, to the 50, 60 students we have now," Lang explained.

According to Childress, the CURE-Purdue Program aims to provide guidance and support to instructors across a range of disciplines to develop, implement, and assess a CURE.

"Course-based undergraduate research experiences are an attractive model for engaging more students in authentic research because they can operate at a larger scale than the traditional apprenticeship model of undergraduate research experiences (or UREs)," noted Amy Childress, Director of the Office of Undergraduate Research (OUR).

"This increases the overall capacity for undergraduate research. In addition, CUREs are more inclusive than the apprenticeship model. Often, when this model is used, the researcher has to be more selective in the students chosen to join UREs. It also requires students to know how to find such opportunities," she explained. "The current body of evidence suggests that CUREs yield many positive outcomes for both faculty and students."

According to Childress, the CURE-Purdue Program aims to provide guidance and support to instructors across a range of disciplines to develop, implement, and assess a CURE.

"The program will provide a larger number of students with access to this valuable experiential learning opportunity, in which they will learn and use the practices of their disciplines and be a part of the discovery of new knowledge," Childress added.



Beyond the Traditional Lab Experience

Wright and Kincaid noted that they have plans to pursue careers in academia; both are planning to attend graduate school after they finish their undergraduate degrees in physics at Purdue. For both, their involvement in Lang's CURE has enabled them to hone their data analysis skills and has provided them with important research presentation skills.

Kincaid said these aspects of her experience have helped her become a better researcher, early on in her research career. She also pointed out that she is genuinely inspired by the data analysis research she is involved with in the XENON1T experiment.

"Sometimes I will think to myself, 'Wow, we really are working on this unique, global project.' Even if we are not finding dark matter, it just feels like a big thing to be a part of it. That is really interesting for me, and it feels important," she noted.

Lang affirmed Kincaid's take on the undergraduate students working on the XENON1T experiment.



Ī₽

ma

в

"For the students involved in the CURE, they are analyzing data at the forefront," he said. "The data that we are analyzing is not only unique, in terms of that it comes from the most sensitive dark matter experiment, but also the sheer volume of it, at more than a Petabyte a year. In physics, there are not too many datasets similar in size. It really is 'big data' because of this size and complexity." - Rafael Lang

12

m=pV

;=umq

.....

mg

tg²φ

dβ

14



A CURE FOR SCALING UNDERGRADUATE RESEARCH CONTINUED FROM PAGE 34

As in many research projects, researchers regularly come together to share and compare work, and the structure of the CURE—with its larger scale and ability to involve a number of undergraduate research teams and graduate research assistants—facilitates this conversational collaboration.

Wright noted that, for him, this is one of the most important parts of his undergraduate research experience.

"The feedback is invaluable. Because sometimes you are going down the

wrong path, and one of the grad students or Rafael will say, 'Maybe you should be looking at this a little more.' And then you can make that switch and start working on that," he explained. "You also get feedback from people that you don't meet with regularly, like grad students you are not paired with; they will also have good ideas for you, to help you learn."

Through his specific CURE, Lang's students analyze XENON1T experiment data that have not been examined previously by the XENON collaboration, opening the possibility for new discoveries.

"It is always funny to call it a dark matter detector because we have not detected dark matter yet, but we're trying to; we are using data from that detector, analyzing it, with the undergraduates here at Purdue," he explained. "In different semesters, students see different things, they are trying to find different kinds of dark matter, or trying to understand different aspects of the detector. It is real research; many of the same things that researchers do in an advanced lab, (e.g., learn presentation skills, statistical analysis, big data). Often researchers do all that in a lab but, the difference here, in this course, is that we very honestly do not know at the beginning of the semester what the outcome is going to be at the end of the semester. So it is research, in the sense that it is completely open as to what the results are going to be. We don't know."



LEARN MORE ABOUT

PURDUE OUR

visit www.purdue.edu/undergrad-research

XENON1T visit www.xenon1t.org.



36

ADVANCING DIGITAL HUMANITIES AT PURDUE

By Teresa Koltzenburg

Developing and advancing Purdue University's Digital Humanities (DH) initiative, a growing area in digital scholarship, comprises most of Assistant Professor of Digital Humanities Matt Hannah's work at Purdue. Since he arrived at Purdue University Libraries and School of Information Studies in March 2018, he has been busily laying the foundation for an ongoing and robust discussion about DH and digital scholarship across campus.

He has designed and delivers a wide range of DH workshops, teaches courses, leads and contributes to digital scholarship

projects and efforts on campus and beyond, and has established the Digital Humanities Studio in the Humanities, Social Sciences, and Education (HSSE) Library. In addition, through the support of the Integrative Data Science Initiative (IDSI) at Purdue, Hannah and his colleague, Venetria Patton (Professor of English and African-American Studies in the Purdue College of Liberal Arts), have been able to award nine members of the CLA faculty small seed grants to redesign existing humanities or social sciences courses to include digital scholarship. In 2018, their

proposal "Engaging in the Digital Humanities" was awarded funds as part of Purdue's "first investment towards achieving the goals of the Integrative Data Science Initiative."

Through Purdue University Libraries and School of Information Studies, Hannah has also been diligently pursuing the development of DH certificate programs for both Purdue graduate and undergraduate students. The idea behind the certificate program is that a student will be able to complete a series of digital humanities and/or digital scholarship courses to earn a DH credential while studying at Purdue.

"Even though at Purdue, humanities disciplines are primarily situated within the College of Liberal Arts, one of the great things about being in the Purdue University Libraries and School of Information Studies is that we occupy both physical and figurative spaces here that can, and do, foster and facilitate collaboration between STEMfocused areas and liberal arts. Because of this, we are wellpositioned at Purdue to be a hub of digital scholarship and Digital Humanities, or what I define as the combining of digital technology and methodology with traditional humanities disciplines, or, in other words, adding quantitative methodologies to qualitative fields of study," Hannah explained. "In Purdue University Libraries and School of Information Studies, we can bring campus units together in interesting ways."





Shaad Hussain (left) and Bridget Arnold (right), a computer science major and an anthropology major, respectively, play an intense game of Boulder Dash in Hannah's Honors 399: Dead Media class. Hussain said the course is teaching him about how media technology has evolved.

At such a STEM-focused institution that aims to be a global data science leader, Hannah has anticipated and fielded the question about how students, once they get out into the real world, will apply DH skills.

"I certainly understand this is an important question, but, also, when I talk with students, I tell them I think there is value in resisting the 'how will you apply this?' discussion, at least to some degree. Humanities have intrinsic value for our society beyond just gaining employment. But what I do, when teaching DH and about some particular DH methodology or tool, is ask students and faculty, 'will this actually help your research?' If it won't, then don't use it. Not everything is a nail if you have a hammer," he said. "On the other hand, there has been study after study that show tech companies also want people with liberal arts degrees, who are creative and who can think outside the box. Often, when you combine the creative thinking and critical thinking you get from the liberal arts with some programming skills, that's a very powerful combination."

Hannah also noted the DH opportunities he is developing at Purdue are meeting a need, for humanities students in particular, that has not been addressed in a comprehensive way here previously.

"One of the things we've already seen is that students in humanities and social sciences are hungry for opportunities to learn digital skills in a non-STEM context, in a very humanitiesfocused way," Hannah said. "These skills can translate into all manner of jobs outside of academia. So, for example, undergraduates who want to get English degrees can learn some programming skills, and combine those seemingly different backgrounds into very productive careers in many fields." One of the things we've already seen is that students in humanities and social sciences are hungry for opportunities to learn digital skills in a non-STEM context, in a very humanities-focused way.

- Matt Hannah

According to Hannah, the DH certificate program will first be available to graduate students. The program for undergraduates will be structured similarly, but will not have as rigorous requirements. He said that the program curriculum is comprised of two required courses and two electives.

ADVANCING DIGITAL HUMANITIES AT PURDUE CONTINUED FROM PAGE 40

"All students will be required to take 'Introduction to Digital Humanities, the 'gateway' course, which I have been teaching since Spring 2019. This course covers the major tools and methods prominent in the study of digital humanities, and in it, students also read a lot about the theory and cultural criticism behind those methodologies. For example, students may be thinking about maps and cartography in a theoretical way, but then they are also learning to use a tool like ARC GIS to create maps," Hannah said. "The goal behind this course is to get students up to speed using certain tools, but also to get them thinking critically about all the tools in the tool box, so to speak, and the affordances and limitations of each one of those tools."

Hannah said the other required course in the DH certificate is a capstone experience. "This is a semester-long mentor-mentee relationship with me, and during it, they will build a digital project," he explained. "Usually, they will take the digital project they started in the introductory course, and they will develop it into a more robust, publicfacing project."

Through such digital scholarship initiatives, Hannah and his colleagues (from **Purdue University Libraries** and School of Information Studies and other colleges) are aiming to position Purdue University as a leader in Digital Humanities. This will provide a unique learning environment, as humanities have always been an important part of the learning equation for students. With DH, humanities can now be paired with other interests in computer science, engineering, data science, and more.

"One way I am interested in thinking about digital humanities is that the core of what we are doing is still humanities. Yes, through the 'digital' part of DH, we're applying quantitative elements and digital methodologies to the study, research, and teaching of the humanities. But we're also applying humanistic forms of inquiry to the digital, whether in critical data studies, science and technology studies, or data science. We are thinking about how we are incorporating data into our lives, and we're thinking about the relationship of humanities to technology. At a place like Purdue, that can be very interesting."



Hannah engages with students in his Honors 399: Dead Media course, in which they have the opportunity to experience video games from the late 20th century. The course covers how media and technology have evolved.



Libraries and School of Information Studies Libraries Administration 504 West State Street West Lafayette, IN 47907-2058

Nonprofit Organization U.S. Postage PAID Purdue University

An equal access/equal opportunity university Produced by Purdue University Libraries and School of Information Studies