Student scientists, dusty data, and dirty discoveries

November 4 2015

Modern technology has blurred the boundaries between place, time zone, and people—between students learning details, and scientists leading discoveries.

Case in point: A revolutionary virtual class brought together undergraduate and graduate students at the University of Arizona (UA) and Western Michigan University (WMU) in the spring of 2014.

With an innovative academic curriculum culminating in a student-authored research paper published today in *PLOS ONE*, the students and professors of the class, titled 'Ecoinformatics' have demonstrated that there are no boundaries remaining for education, collaboration and scientific discovery.

Redefining Collaborative Learning

"From the beginning it was our goal that the students would write a paper. I don't think they believed us at first." Rachel Gallery is a UA assistant professor in the School of Natural Resources and the Environment in the College of Agriculture and Life Sciences.

Gallery and Kathryn Docherty, an assistant professor of biological sciences at WMU, pioneered Ecoinformatics, a class that integrated remote collaboration, data-sharing, and peer-review publication into a new form of learning that goes well beyond customary college

coursework.

The class convened for lectures via videoconference.

"It was surprising how easy it was to treat it like a normal classroom," Gallery remarked. "In some ways it was more helpful than a standard lecture, much more effective because we spent less time lecturing and more time having discussions."

Martha Gebhardt, a UA doctoral candidate and Ecoinformatics student, agreed.

"Virtual lectures facilitated class discussions," she said. "Both UA and Michigan students could immediately see what was being discussed and add their thoughts and input."

The class combined undergraduate and graduate students from a variety of educational and intellectual backgrounds, including entomology, soil sciences, environmental sciences and informatics.

The mixture provided diversity that Gebhardt found inspiring: "Everyone had something unique and beneficial to bring to the table. Concepts were often discussed or presented in ways I have never considered before."

"Some of the background information and principles we employ as graduate students can be taken for granted, and having undergraduate input was important for identifying some of our working assumptions," agreed Hannah Borton, a doctoral candidate in biological sciences at WMU who took Ecoinformatics. "I gained a valuable appreciation of other students as colleagues.

The two young professors asked the students to outline what they wanted

to learn and what they could teach others, and developed the course agenda from that information. Students were assigned teams to focus on topics aligned with their interests.

"Some wanted to learn how to code, some wanted to learn more about literature reviews and writing manuscripts, some were interested in informatics, the data-wrangling," Gallery said. "The students were incredibly engaged."

"Not only did we get to choose an area we were most interested in pursuing, but we had other students in the group equally interested in the topic," Gebhardt said. "During class time each group would present progress they made the previous week, often informing future directions."

"It was in many ways a professional development course," reflected Gallery. "We taught students concepts including how to write a manuscript, who deserves authorship and co-authorship, and how do you allocate those responsibilities?"

But before they could write a paper, the students had to learn to analyze large-scale datasets. Modern technology has fueled the [big data](#) revolution with new and more powerful resource tools generating huge amounts of data, often more data than scientists have time or resources to study.

The massive volumes of unanalyzed data thus are funneled into so-called big data repositories, science centers that store and catalog the datasets with the hope, and intent, that they will someday be used to pioneer new discoveries.

The result is 'open-access data,' free for anyone scientifically inclined to mine for answers to questions that oftentimes haven't even been asked,

leading to valuable new knowledge.

Researchers everywhere are talking about using open access data, Gallery said. "So we tried training students on big data questions, and had so much success that it resulted in a manuscript."

That manuscript was accepted, and published today, by *PLOS ONE*, a prestigious open-access, peer-review journal.

Sifting Through Soil Secrets

Inspired by access to untouched scientific data and unbarred restrictions on what they could do with it, the students of Ecoinformatics transformed into eager scientists.

They weren't memorizing seemingly senseless details—they were mulling over scientific discoveries. They used social media and online discussion forums to collaborate, discuss, and strategize. They dove headfirst into the data, ideas sparking.

"These students took us in directions we weren't even thinking about in terms of questions that you could ask with these data," Gallery said.

The students leveraged previously unanalyzed pilot data from the National Ecological Observation Network, or NEON, an observation system designed to enable researchers to examine ecological variation over time, on a continental scale.

The class developed the scheme of analyzing the effects of geography and temperature on soil bacteria communities in four different biomes. Biomes are ecological zones characterized by ability to support distinct communities of life forms.

The students selected datasets collected from biomes in Utah, Hawaii, Alaska, and Florida, and began evaluating seasonal variation of terrestrial vegetation and comparing peak growing season values across the four biomes.

To profile the microbial communities, the students used bacterial DNA and lipids, or oils, produced by bacteria and fungi in the soils, which provide an estimate of their growth.

"It was special," Gallery said. "A student-driven concept."

Docherty and Gallery both had been involved with the initial microbial sampling by NEON and were familiar with the data, perfectly positioning them to guide their students' understanding.

To securely store, share, and analyze the massive volumes of data, the class turned to the iPlant Collaborative, a National Science Foundation-funded biotechnology project that provides computational resources for big data storage, analysis, and sharing.

The class, and its research, education, and publication outcomes, would not have been possible without the iPlant Collaborative, Gallery said. She and Docherty leveraged online educational tools and services provided by iPlant to help the students learn how to analyze big data.

"We used iPlant for all the data sharing, so everybody could access the data," Gallery continued. "And the students used the iPlant environment to send messages as they worked their way through the data."

The students even developed their own idea of creating YouTube videos to help teach each other the various skills needed for success in the course, including how to use iPlant's services.

From their research and correspondence, the students determined that key properties, such as soil temperature, soil chemistry and vegetation, could explain most variation in soil bacteria across the four biomes. The research data from the course are available through the iPlant Collaborative, and are stored with corresponding metadata in a public data repository by the National Center for Biotechnology Information, an initiative of the National Institutes of Health.

And from these data, the students of Ecoinformatics co-wrote what for many was their first scientific publication.

"It is first and foremost a research paper," Gallery noted. "But we also talked about the usefulness of this approach for project-based learning."

Lessons Never Learned from Lectures

"As a graduate student, having a publication really demonstrates commitment to your research, and ability to perform," said UA doctoral candidate Noelle Espinosa.

"This course provided so much more than most courses. We were challenged to think and work as a collaborative group, to ask big questions and grapple with a big dataset. What I picked up from my peers will be invaluable for my future."

"Skills practiced in this [class](#) are integral to feeling prepared to be a professional scientist," agreed Borton.

"Oftentimes researchers feel limited to collaborating just with local researchers, but with these large datasets and today's communication and data-sharing systems, that is no longer a limitation," Docherty noted.

As the [students](#) wrote in their publication: "Increased availability of

public datasets ... and improved data sharing platforms, (e.g., the iPlant Collaborative) is becoming representative of the future of ecological research."

More information: *PLOS ONE*, dx.plos.org/10.1371/journal.pone.0135352

Provided by University of Arizona

Citation: Student scientists, dusty data, and dirty discoveries (2015, November 4) retrieved 27 April 2024 from https://phys.org/news/2015-11-student-scientists-dusty-dirty-discoveries.html