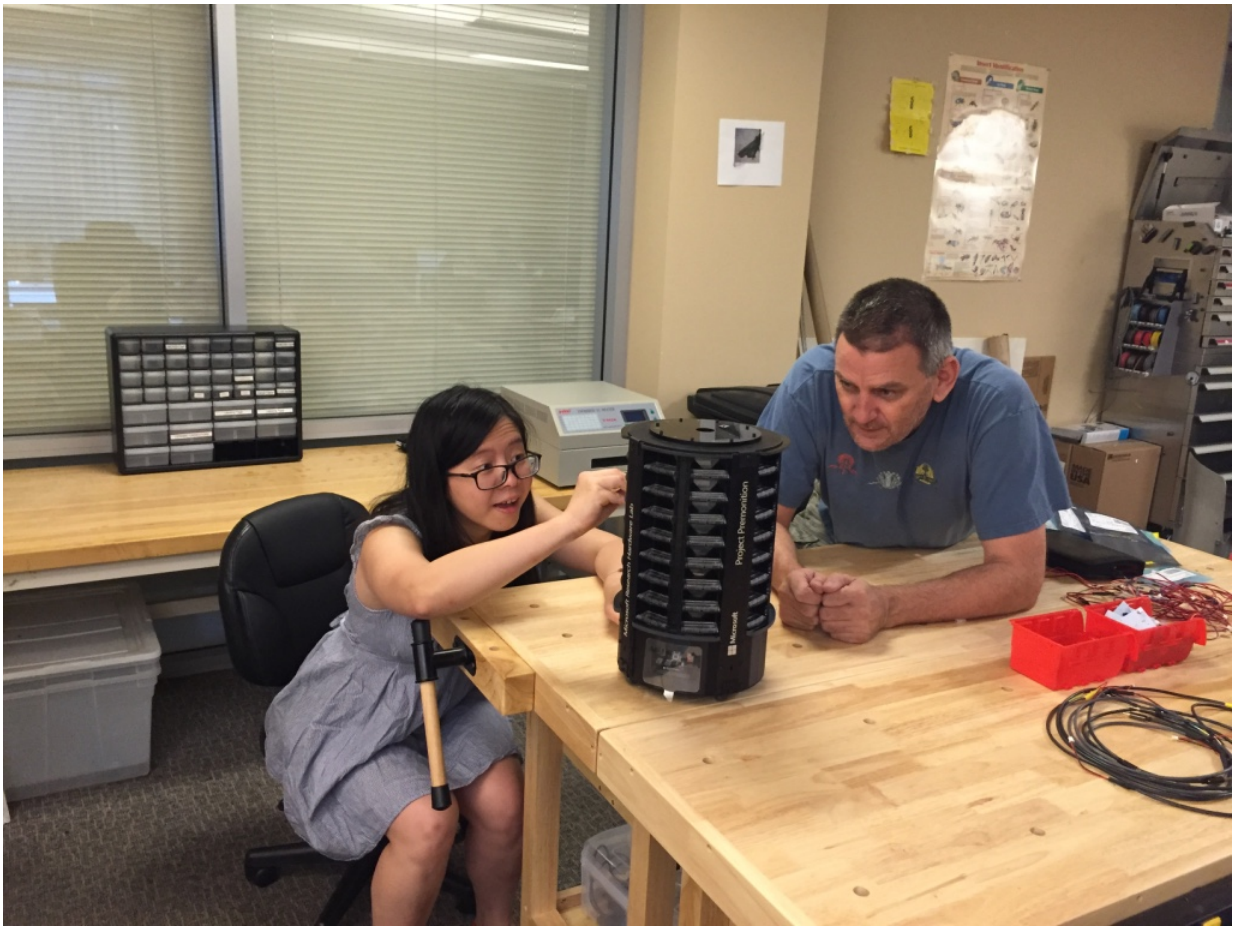
# Big data for little creatures

October 11 2016, by Sarah Nightingale




Doctoral student Yan Zhu (left) and Eamonn Keogh, professor of computer science and engineering, adjust a mosquito trap designed by researchers at UCR and Microsoft. Credit: UC Riverside

While millions of people are fighting Zika with bug spray and long

pants, researchers at the University of California, Riverside are using another tool—big data. By collecting and analyzing large datasets, they can track the mosquitoes that spread the disease in real time, intervene quickly, and help governments plan for future outbreaks.

This is just one example of research combining computer science and entomology that will be the focus of UCR's new graduate program in computational entomology. Led by Eamonn Keogh, a professor of computer science and engineering in UCR's Bourns College of Engineering, a multi-disciplinary team of researchers has received $3 million from the National Science Foundation Research Traineeship (NRT) program to prepare the next generation of scientists and engineers who will learn how to exploit the power of big data to understand insects. The program, the first of its kind worldwide, will serve as a replicable education and training model for other institutions with an interest in developing computational entomology programs.

There are at least 950,000 living insect species in the world. They spread diseases like Zika virus and malaria; they feast on crops, causing billions of dollars in food damage each year; and they pollinate wild and agricultural plant communities. Despite their enormous diversity and significance, few scientists have tapped into the power of computer science to classify insects and explore their behaviors.

For the past five years, Keogh has been doing just that. The researcher and his graduate students have been developing low-cost, wireless insect sensors that classify species with up to 99.9 percent accuracy and generate masses of data that can be incorporated into classification algorithms. In about three years, and with dozens of sensors running continuously, the team collected tens of millions of data points—more than all previous work in the field combined.

By counting and classifying insects on this scale, scientists can help

farmers determine precisely when to apply pesticides, and help public health officials stop the spread of insect-borne diseases. Many other areas of entomology would benefit from such in-depth analyses, Keogh said.

To develop the new program, Keogh will work with UCR faculty members Erin Wilson Rankin and Anupama Dahanukar, both assistant professors of entomology; Daniel Jeske, professor of statistics; and Christian Shelton, professor of computer science and engineering. Called the 'NRT in Integrated Computational Entomology' (NICE), the program will launch next summer and fund at least 80 graduate students enrolled in engineering or the life sciences. It will begin with a month-long introductory boot camp and feature weekly seminars, field trips, and community service projects. There will be a strong emphasis on collaboration across disciplines—with biological sciences students gaining a foundation in computing techniques and engineers an understanding of critical entomological and ecological issues.

Keogh said students will pursue their own research interests. Some might continue his work on sensors, developing real-time tracking devices that send text messages to farmers letting them know when a harmful insect is multiplying. Some may use video analytics to study insect behavior and how to control it. Some will sequence mosquito DNA to pinpoint genes that determine prey preference.

The common thread, Keogh said, is the power of big data to expand our understanding of insects, improve human health, and alleviate food waste.

"Challenges in entomology and ecology generate enormous amounts of data, and fully exploiting it calls for experts whose knowledge spans two disparate fields. This training program will bridge those fields, creating endless research possibilities and a new way to address some of the most

critical challenges of our time," Keogh said.

Provided by University of California - Riverside