

BIOMEDICAL DATA SCIENCE
11/3/22
MSOB 303X
1:30 PM - 2:50 PM

David Van Valen is an Assistant Professor in the Division of Biology and Bioengineering at Caltech. Engineering and Biomedical Engineering

TITLE: EVERYTHING AS CODE

ABSTRACT:

Biological systems are difficult to study because they consist of tens of thousands of parts, vary in space and time, and their fundamental unit—the cell—displays remarkable variation in its behavior. These challenges have spurred the development of genomics and imaging technologies over the past 30 years that have revolutionized our ability to capture information about biological systems in the form of images. Excitingly, these advances are poised to place the microscope back at the center of the modern biologist’s toolkit. Because we can now access temporal, spatial, and “parts list” variation via imaging, images have the potential to be a standard data type for biology.

For this vision to become reality, biology needs a new data infrastructure. Imaging methods are of little use if it is too difficult to convert the resulting data into quantitative, interpretable information. New deep learning methods are proving to be essential to reliable interpretation of imaging data. These methods differ from conventional algorithms in that they learn how to perform tasks from labeled data; they have demonstrated immense promise, but they are challenging to use in practice. The expansive training data required to power them are sorely lacking, as are easy-to-use software tools for creating and deploying new models. Solving these challenges through open software is a key goal of the Van Valen lab. In this talk, I describe DeepCell, a collection of software tools that meet the data, model, and deployment challenges associated with deep learning. These include tools for distributed labeling of biological imaging data, a collection of modern deep learning architectures tailored for biological image analysis tasks, and cloud-native software for making deep learning methods accessible to the broader life science community. I discuss how we have used DeepCell to label large-scale imaging datasets to power deep learning methods that achieve human level performance and enable new experimental designs for imaging-based experiments.

Website: <https://vanvalen.caltech.edu>

Zoom info: [https://stanford.zoom.us/j/92874055477?](https://stanford.zoom.us/j/92874055477?pwd=aThZNmpmNEQ1L2FjV0E5ZXF5S0R1UT09&from=addon)
pwd=aThZNmpmNEQ1L2FjV0E5ZXF5S0R1UT09&from=addon
Password: 705300