



# CellProfiler: Open-source, high-throughput image analysis software

Anne Carpenter and the members of the Imaging Platform at the Broad Institute of Harvard and MIT



**Overview & Introduction:** Microscopy images contain rich information about the state of cells and organisms and are an important part of experiments to address a multitude of basic biological questions and world health problems. Our laboratory works with dozens of collaborators around the world to design and execute high-content screens using chemical libraries or genetic perturbations, in order to identify the causes and potential cures of disease. In this poster, we overview this work. For instance, we are working to identify novel drugs for various forms of leukemia and infectious diseases, regulators of DNA damage, distinctions between human isoforms of cancer-relevant histone deacetylase proteins, mechanisms of hepatotoxicity, and diagnostics for bipolar disorder and schizophrenia. Biologists are developing model systems that are more and more physiologically relevant, yet still compatible with automated instrumentation. Such systems include co-culturing two different cell types to better mimic functional tissue and culturing whole organisms such as *Caenorhabditis elegans* to study entire organ systems. Machine-learning approaches, guided by a biologists' intuition, have been particularly successful for measuring subtle phenotypes in these increasingly complex model systems.

The image analysis algorithms and data mining approaches we develop are freely available in the biologist-friendly open-source software, CellProfiler, for small and large-scale experiments: [www.cellprofiler.org](http://www.cellprofiler.org)

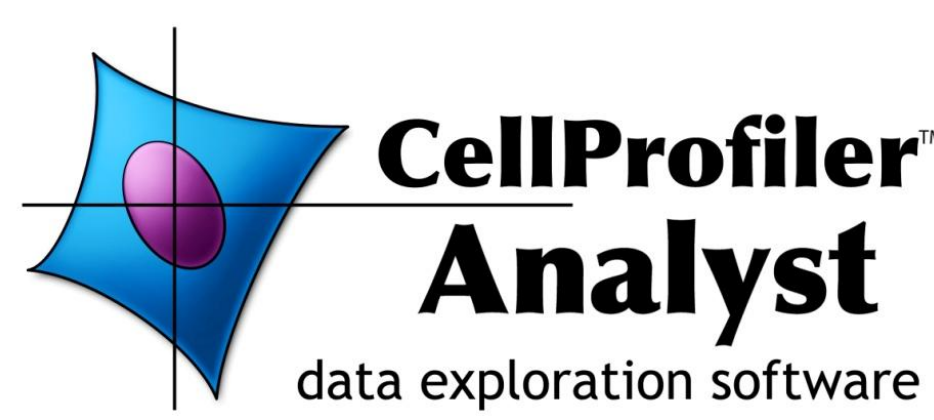
## Method

S



**CellProfiler** cell image analysis software is designed for biologists without training in computer vision or programming to quantitatively measure phenotypes from thousands of images automatically.

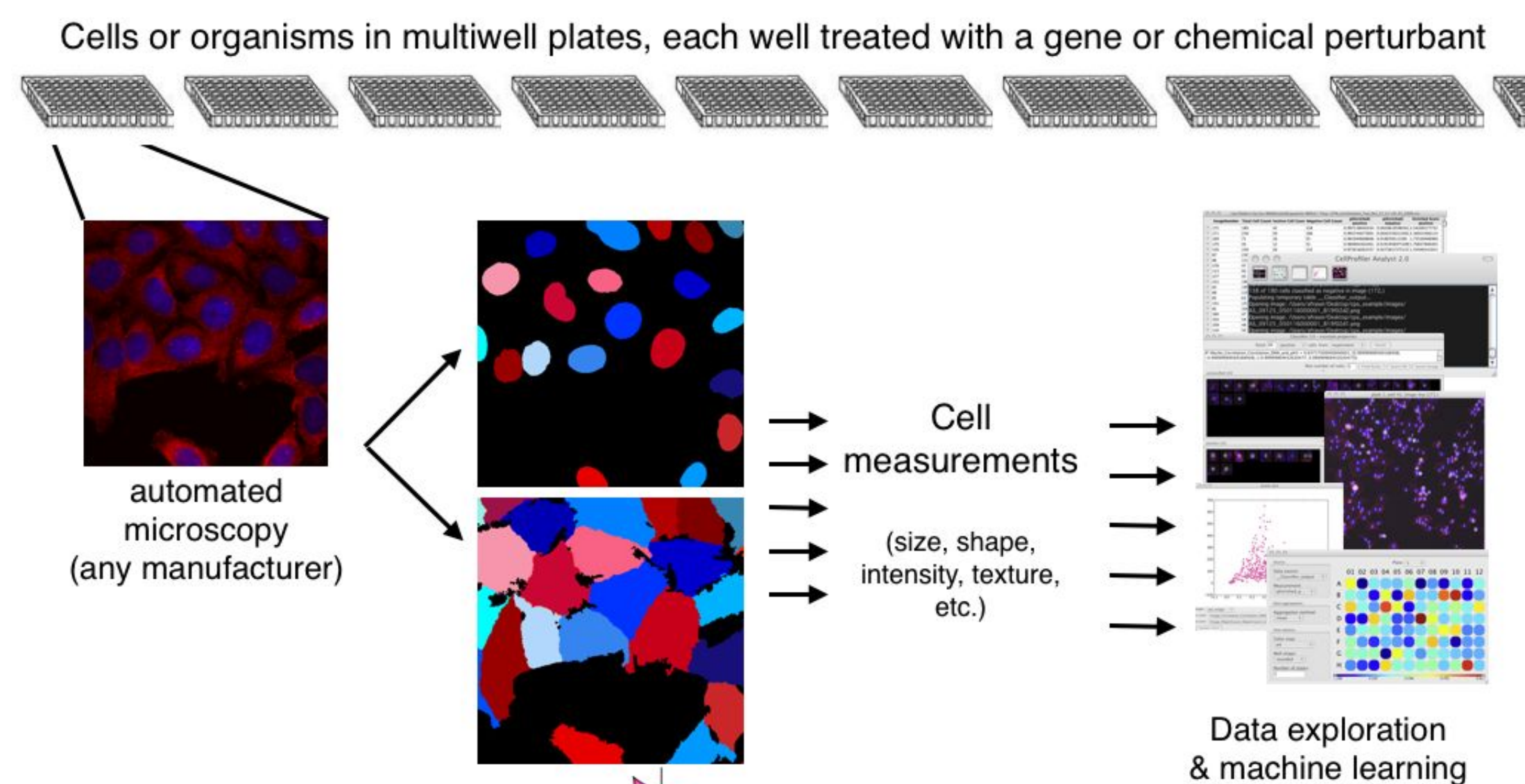
Using the flexible modules within CellProfiler, researchers set up automated image analysis pipelines to identify a wide variety of biological “objects”, including yeast colonies, neurons, cell co-cultures, and whole organisms such as *C. elegans* worms. Multiple features of each object and associated sub-compartment are then measured, including size, shape, intensity and texture of each color channel in the original image.



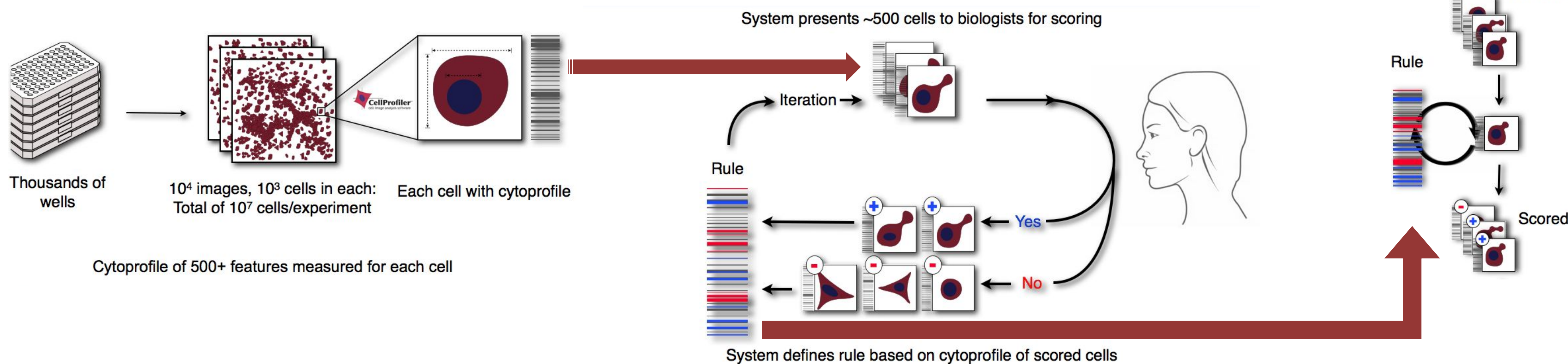
**CellProfiler Analyst** allows interactive exploration and analysis of data, particularly for high-throughput, image-based experiments. Included is a machine learning system which can be trained to recognize complicated and subtle phenotypes.

When an assay is complex, it is often not clear which features are suited for classifying the cells or other biological objects as “positive” or “negative” for the phenotype of interest. CellProfiler Analyst provides a powerful machine learning tool to enable classification based on multiple morphological features of each object, as measured by CellProfiler. CellProfiler Analyst also contains data visualization tools to help researchers analyze the output of large experiments.

## Image analysis algorithms identify and measure cells in high-throughput



## Machine learning scores complex phenotypes



## Results: Phenotypes measured using CellProfiler and CellProfiler Analyst

