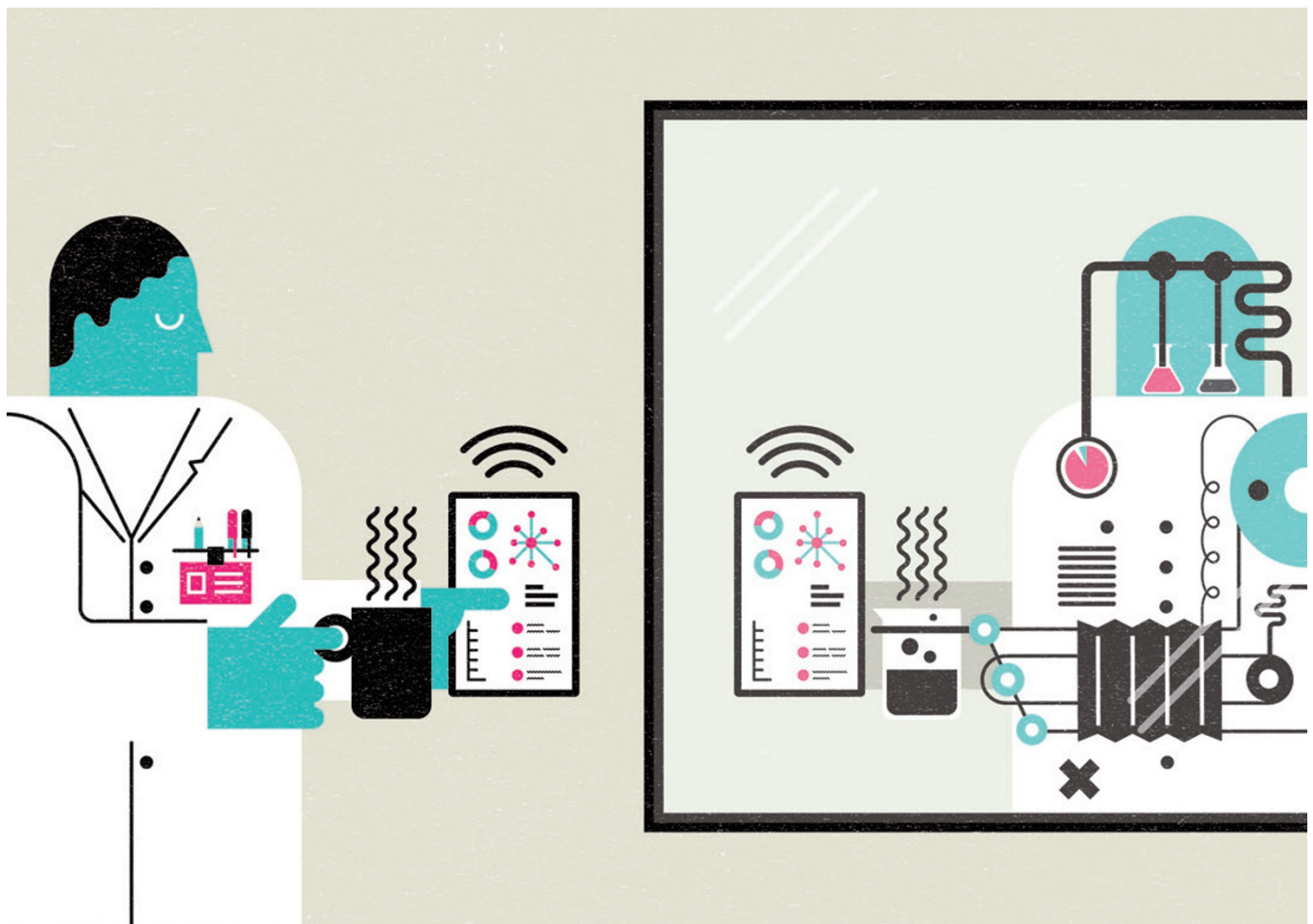


TOOLBOX

THE AUTOMATED LAB

Start-up firms say robotics and software that autonomously record every detail of an experiment can transform the efficiency and reliability of research.

ILLUSTRATION BY THE PROJECT TWINS



BY ERIKA CHECK HAYDEN

Max Hodak has spent much of his academic career fixing the ways that scientists collect data. As a biomedical engineering student at Duke University in Durham, North Carolina, it frustrated him that his laboratory recorded its experiments in paper notebooks, leaving researchers to scour through the pages to find relevant data. So in 2008, he indexed all the notebook data on a computer

and wrote a program to allow users to query it. “People were saying, ‘Why are you wasting your time? That’s not going to lead to publication,’” he recalls. But a year-and-a-half later, he returned to the lab from a stint in Silicon Valley to find that many of those earlier sceptics were now using his system. To Hodak, it was a sign that he should pursue his quest for efficiency in the lab. “I was always more interested in finding ways to do analysis more efficiently than in doing the actual analysis,” he says.

Today, a warehouse in California is the living embodiment of Hodak’s dream to build an automated lab that conducts experiments and records the results, or what he calls a “biology data centre”. His company, Transcriptic, founded in 2012, is the first of a crop of start-ups of this ilk, all with a similar claim: that advances in software and robotics will help to free researchers from manual drudgery, make their data easier to store and query, and ultimately lead to cheaper, more ▶

► efficient and more reproducible science.

Transcriptic and another California firm called Emerald Therapeutics are pinning their hopes on offering scientists control of a wet lab by remote computer. Many big biology labs already have automated machines to sequence or copy pieces of DNA. But these companies want to bring automation to other routine experiments, such as moving and separating proteins or fragments of DNA through a gel. They offer the capabilities to labs big and small.

In Transcriptic's approach, customers first program an experiment using the company's application programming interface (API), which translates each step of an experimental protocol into machine-readable code. The customer's orders — and any physical samples — arrive at Transcriptic's warehouse in Menlo Park, and the experiment is carried out in a Plexiglas-enclosed station that contains a benchtop full of instruments guided by a computer, which receives the work order and runs the assemblage of machines. A robot on a gantry runs the length of the station, transferring plates from machine to machine — apparatus for the polymerase chain reaction (which amplifies DNA), plate readers, liquid handlers, a freezer and an incubator — to carry out the experiment. Users receive data from each step of the experiment in real time.

Customers say that this frees them up to spend their time on science, rather than on grunt work. For example, synthetic biologist Justin Siegel heads a lab at the University of California, Davis, that designs, builds and tests new enzymes. Transcriptic now does the lab's molecular biology, liberating his students from the one-third of their time that used to be spent on copying and mutating DNA fragments (cloning and mutagenesis, respectively). Siegel credits Transcriptic with building a biosensor that detects the chemical profile of olive oil, for which his students won the grand prize at the 2014 International Genetically Engineered Machines competition on 3 November. "It's made us more efficient and a little bolder," he says. "Instead of making just ten designs, they want to try a couple extra. They'll go a little bit farther out on a limb because all of a sudden they don't have to physically build the stuff."

EXPERIMENTS TO ORDER

Emerald is testing what it calls the Emerald Cloud Laboratory, which the company says will be a one-stop online shop through which customers order experiments, analyse data and collaborate with others. Starting early in 2015, beta users will be able to order from a list of 40 common lab protocols, such as western blots for protein analysis, or high-performance liquid chromatography for separating components of a mixture. When an order comes in, a human operator will set up the experiment on one of the company's automated workstations

at its lab in Menlo Park. Operators transfer sample plates from machine to machine to carry out the steps of the experiment, and the customer gets data back through the Emerald Cloud Laboratory. There, customers can analyse the results using text functions encoded in the Wolfram programming language. Users can review everything, from the controls and machine settings of the instruments used, to results from past experiments using the same reagents — even experiments from others who have granted permission for their data to be openly accessible. "Everything we do is built by scientists for scientists," says Emerald co-founder Brian Frezza.

In one sense, these models are similar to conventional contract-research organizations, but the automated systems and data collection offer scientists much finer detail and control over experimental design. One might think that the complicated equipment makes experiments more expensive to do. But Hodak says that Transcriptic offers protocols such as cloning and mutagenesis at about the same cost at which they could be run in an academic lab — or for less — and at about half the price offered by conventional research outsourcing firms, he says. That is partly because its stations can run without a human operator, and also because the firm makes a lot of its own hardware. When Hodak first went looking for an automated freezer, the cheapest he could find was US\$400,000. So he hired mechanical engineers to build one for \$40,000, and used the same basic design for automated incubators and refrigerators.

Although these labs are powerful research aids, remote users are necessarily limited to a standard set of experiments and instruments, notes Roger Chen, an associate at the investment firm O'Reilly AlphaTech Ventures in San Francisco, California. "I have a hard time believing that a centralized automated lab will give you the freedom and flexibility to experiment with all the parameters you need to do some innovation."

Chen has therefore invested in Riffyn, another fledgling start-up, in Oakland, California. It wants to adapt individual research labs so that scientists can use automated data collection for their own custom experiments, albeit without robotic control. The firm (which will not launch until late 2015 at the earliest) is building a cloud-computing

➔ **NATURE.COM**
For more on scientific software, apps and online tools, visit: nature.com/toolbox

software platform integrated with devices that stream data from lab equipment. According to co-founder Timothy Gardner, the software

will allow users to design workflows, analyse experimental data held on remote servers and thus change parameters (such as the temperatures or pressures at which an instrument operates) in response to the analysis of an experiment's performance. "We're trying to solve the problem of how you bring McDonald's-like efficiency to scientists without shackling them to McDonald's-like recipes," he says.

COMMUNICATION IS KEY

Current scientific instruments take instructions and record data in a variety of formats. Harmonizing their software will be an even harder technical challenge than building robotically controlled equipment, warns Frezza. Gardner acknowledges this, but says that momentum for a common, open set of standards and software is building. Meanwhile, he hopes that Riffyn's software will enable lab devices to talk to each other — and to scientists — more smoothly. The concept would not have been possible even a few years ago; only last year did computer giant Apple, for instance, release iBeacon, a system that enables nearby compatible devices to communicate with each other.

Gardner's inspiration comes from his time running research operations at synthetic-biology pioneer Amyris Biotechnologies, in nearby Emeryville, which genetically engineers yeast to produce biofuels and specialty chemicals. In the early days it struggled to scale up basic processes to industrial quantities, because random variation — noise — in each step led to irreproducible results. The company started to analyse data from each step of the process to spot the weak points, and made dramatic improvements. Gardner says that research labs need that same kind of reproducibility. "We need to bring the pursuit of precision reliability to the academic world," agrees Douglas Crawford, associate director of the California Institute for Quantitative Biosciences (QB3), and one of Gardner's strongest supporters.

Ultimately, Gardner and other backers of automated labs say that the immediate payoff of their work might be to promote a general movement to boost the overall quality of research. Tools that make it easy for scientists to monitor and record every aspect of their experiment, they say, might help to deal with what some argue is a 'reproducibility crisis' in research — the sense that many experiments are too sloppily done, or that methods and data are recorded too imprecisely, for others to easily reproduce findings.

Crawford thinks that changing this will require a cultural shift as much as a technical one. Gardner agrees, but hopes that companies such as his will remove one the roadblocks: "I don't think you can get the cultural and educational changes to stick if you don't have tools that make it easy," he says. ■