Emerald Cloud Lab's facility for running experiments will operate 24 hours a day and allows users to control instruments remotely.

EMERALD CLOUD LAB

AUTOMATION

An operating system for the biology lab

Laboratory-automation start-ups are borrowing a page from the software industry.

BY MICHAEL SEGAL

"Their approach is beautifully simple," says Dhash Shrivathsa, founder and chief executive of Radix Labs. "They put into a room all the equipment needed to process and sequence the genetic material in a vial of human blood." He's describing a leading liquid-biopsy company that has raised hundreds of millions of dollars on the promise of early disease detection on the basis of drawing blood. Its machines are highly automated and engineered to run seamlessly — and can be painstaking to reprogram when the design of an experiment changes. The company was spending months at a time doing so. "It was like switching from one program to another on your computer without an operating system," explains Shrivathsa.

To speed up that process, the company turned to Radix for help. The two-year-old start-up in Cambridge, Massachusetts, has developed a computer language that can be used to encode a customer's biology experiments, a compiler that translates that code into a machine-readable language, and a set of drivers that enables instructions to be understood and executed by the customer's equipment. Radix's technology aims to free biologists from worrying about the details of the machines in their laboratories and how they execute an experiment. This enables — among other things — the rapid redesign of experiments without needing to manually reprogram (or even physically reposition) the equipment involved. "If two different robots in the biology lab are told to do the same thing, we'd like to be able to have our customers use the exact same program on either or both, without any code change or effort on their part," says Shrivathsa.

Automation has been integral to biology labs for decades — for example, in the form of programmable centrifuges and mixers. These instruments ease repetitive tasks and speed up the execution of experiments. But Radix's goal is more ambitious. Just as a software engineer writes code for an idealized and virtual computer, the location and details of which are irrelevant, Radix's technology enables a biologist to plan experiments for a virtual biology lab. The procedures are still carried out using real-world equipment, but the scientist does not need to know its make and model, or how and where it is set up. This represents both a practical and conceptual shift in how biology is done. The hardware layer separating the biologist from his or her subject becomes slightly more transparent.

Virtualization of the biology lab is an ambitious project. "Coming to this as a biologist, my initial instinct was that it wasn't amenable to this treatment," explains Markus Gershater, chief scientific officer at Synthace, a biological-software firm that he co-founded in London in 2011 and a competitor to Radix. But he was attracted by the possibility of making it easier for biologists to perform experiments in which many variables are modified simultaneously, by freeing them from the proliferation of procedures and handling steps that such experiments usually require. Contrary to his expectations, he found that experimental biology is well-suited to virtualization, because its underlying procedures are so similar. "So much of the logic in biology is in moving liquids around," explains Gershater, "whether you're working on malaria or genetically modified products."

Radix, Synthace and other companies that are pursuing the virtualization of biology labs have lofty goals, and have attracted tens of millions of dollars in investment and some well-funded customers. But these companies are also clear that the challenges they face are not just technical. Their industry will need both cultural acceptance and a set of global standards to reach its potential, just as happened with the software and semiconductor industries that have inspired it. But where these will come from is not entirely clear.

RUN IT AGAIN

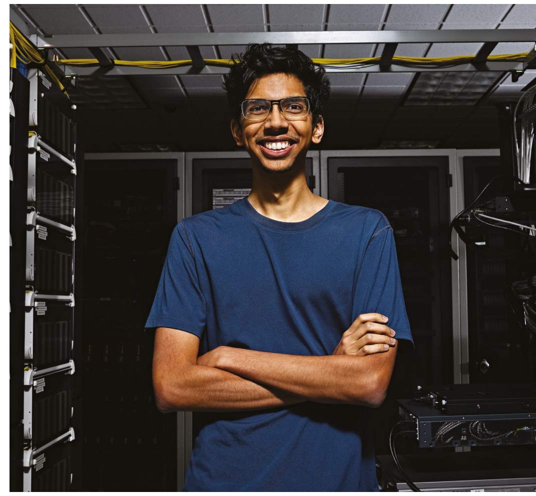
Emerald Cloud Lab in South San Francisco, California, has also developed a computer language and set of drivers that are capable of translating complex experiments into machine-readable instructions. But unlike Radix, which aims to work with the existing infrastructure of its customers — and requires it to write drivers for many kinds of equipment — Emerald owns and maintains the hardware that is needed to run a customer's experiments. After its facility is completed next year, customers will be able to upload a script that encodes their experiment and then simply hit 'Run'. There will be people among Emerald's machines, but they will mostly also be following software-generated instructions — shifting and scanning sample vials and agar plates. "It's a bit like Amazon fulfilment," says Emerald co-founder Brian Frezza.

Emerald's approach, says Frezza, offers a particularly robust version of the reproducibility that every company in the space is pursuing. "Virtualization means you can hand your methods to somebody else, and they can push a button to run the thing exactly again," he explains. When those methods are run on the same lab equipment time after time, uncertainty is reduced even further. This can be profoundly important for some of Emerald's customers, including pharmaceutical companies that are seeking regulatory approvals from the US Food and Drug Administration. Frezza says that researchers can spend months producing dozens of pages of documentation to satisfy government regulations, even for a simple experiment. It's a Herculean effort that could, in principle, be avoided by providing drug regulators with the computer script that was used to run relevant experiments.

Customers can also benefit from the regularity of the data that is produced by Emerald's standardized hardware and software environment. Because each piece of data is linked permanently to the unambiguous script with which it was produced, "it puts this very rigid, searchable ontology on top of everything", explains Frezza. That makes the data easier to analyse. Even failures are useful, because the conditions under which they occurred are perfectly clear.

In addition, the very act of encoding experiments in a programming language can boost reproducibility, and is a key part of the value proposition offered by Synthace and its competitors. Gershater recalls learning, while helping to encode a client's lab processes, that one of its standard procedures was to thaw cells by running them under warm tap water. When pressed on the temperature, they explained that the tap water was 37 °C. "I don't know of a warm tap that's reliably and stably 37 °C," Gershater says. Despite working at a sophisticated company, he notes, the biologists had fashioned a procedure that was potentially irreproducible.

Once an experiment has been encoded, Synthace's software environment, Antha, is able to identify which variables the outcome of the experiment is most likely to be sensitive to — and which probably don't matter at all. For example, by looking at many experiments and contexts, Gershater's team has shown that salt is an unnecessary component of lysogeny broth, a medium that is widely used to culture bacteria. This finding, which simplifies a broad swath of experimental protocols, might never have been made by conventional means. "No one would bother to look at this," Gershater explains.



Dhash Shrivathsa, founder of start-up Radix Labs.

DOUG LEVY AT THE ENGINE

STANDARD OPERATING PROCEDURE

Ginkgo Bioworks, a biotechnology company in Boston, Massachusetts, has built a synthetic-biology foundry — a facility where, with the help of lab-automation firm Strateos of Menlo Park, California, they automate the design and testing of engineered organisms. This includes computational design, DNA synthesis and assessing cell performance. "If you tour our foundry, you'll find a huge variety of different hardware vendors represented," says Barry Canton, the company's co-founder. "They all have their own software interfaces and none of it is standardized." This, he explains, means that software drivers specific to each device need to be developed — a painstaking process.

A set of standards would alleviate this issue, and Canton thinks that companies will eventually deliver them. "It'll be the customer–vendor axis that will drive standardization. It's industry that's the biggest consumer and driving force," he says.

"It's simply a different way of thinking about biological experimentation."

Douglas Densmore, an electrical and computer engineer who is working on automation for synthetic biology at Boston University, suggests instead that academics are better placed to produce the standards that the field needs. He draws a parallel with computer languages. "Typically, computer languages are developed outside of a company, in a research atmosphere," Densmore says. "No one sells Java. No one sells C++. They make something with it. I would hope that biology might be similar some day."

Getting academics on board, however, might not be straightforward. Densmore is quick to identify one reason why: many are not incentivized by reproducibility. "Making it easy for people to replicate an experiment is often not the goal in academia," he says. Some might even prefer to boast that a successful bit of research "was done in a way only my lab could do it". Densmore says that he has actually seen a decline in interest in automation among academics in the past decade. "In the early days, I felt like people were very enthusiastic. Then they realized, 'Oh, this is going to require some discipline on my end.' That's part of the challenge."

Adoption by academia has also been slowed by how the virtualized biology lab manages failure. The technology for identifying and correcting errors in computer code is well established. But systems for dealing with failure in a virtualized lab are not — a particular problem for academics.

Paul Jaschke, a biological engineer at Macquarie University in Sydney, recalls working with Transcriptic, a lab-automation company in Menlo Park, while he was a postdoctoral researcher at Stanford University in California in 2015. (In June, Transcriptic merged with 3Scan, an image-analysis company in San Francisco, California, to form Strateos.) At that time, Transcriptic followed a similar business model to that of Emerald, with a closed hardware environment that was maintained on the company's campus. Although Transcriptic's approach worked well for industry customers who wanted to perform standardized procedures, using the service for untested experiments could be difficult. Jaschke discovered that the feedback that Transcriptic's systems generated automatically after a failed experiment was sometimes not detailed enough to work out what to do next. "I'd need to drive down to Transcriptic's facility to figure out what had gone wrong," he explains.

Gershater thinks that it will take a change in culture for the virtualized biology lab to gain greater acceptance. "There's a lot of rightfully sceptical people," he says. Realizing this change might involve replacing manual verification with consistency checks embedded in software, developing business models that separate design and analysis work from experimental execution, or simply communicating a more precise description of what virtualization can offer.

On this last point, Gershater is abundantly clear: "It's simply a different way of thinking about biological experimentation." ■

Michael Segal *is a writer based in Boston, Massachusetts.*