IEEE
SPECTRUM
FOR THE TECHNOLOGY INSIDER | 03.20



# PHOTOGRAPHY'S NEXT GIANT LEAP


The quantum-dot image sensor will revolutionize digital imaging—and it could be in smartphone cameras within five years **P. 24**

IEEE

IEEE
# SPECTRUM

CONTENTS_03.20



**On the cover**
**Illustration for *IEEE Spectrum* by Anatomy Blue**

SHUNK-KENDER/J. PAUL GETTY TRUST/GETTY RESEARCH INSTITUTE, LOS ANGELES, 2014.R.20


**IEEE SPECTRUM**
(ISSN 0018-9235) is published monthly by The Institute of Electrical and Electronics Engineers, Inc. All rights reserved. © 2020 by The Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, New York, NY 10016-5997, U.S.A. Volume No. 57, Issue No. 3. The editorial content of IEEE Spectrum magazine does not represent official positions of the IEEE or its organizational units. Canadian Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40013087. Return undeliverable Canadian addresses to: Circulation Department, IEEE Spectrum, Box 1051, Fort Erie, ON L2A 6C7. Cable address: ITRIPLEE. Fax: +1 212 419 7570. INTERNET: spectrum@ieee.org. ANNUAL SUBSCRIPTIONS: IEEE Members: $21.40 included in dues. Libraries/institutions: $399. POSTMASTER: Please send address changes to IEEE Spectrum, c/o Coding Department, IEEE Service Center, 445 Hoes Lane, Box 1331, Piscataway, NJ 08855. Periodicals postage paid at New York, NY, and additional mailing offices. Canadian GST #125634188. Printed at 120 Donnelley Dr., Glasgow, KY 42141-1060, U.S.A. IEEE Spectrum circulation is audited by BPA Worldwide. IEEE Spectrum is a member of the Association of Business Information & Media Companies, the Association of Magazine Media, and Association Media & Publishing. IEEE prohibits discrimination, harassment, and bullying. For more information, visit https://www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

BACK STORY

03.20

W. PATRICK MCCRAY

# THE ENGINEER'S PERSPECTIVE

"ENGINEERS OFTEN DON'T LEAVE a strong presence in the historical record." That's how W. Patrick McCray politely describes the frustration he and other historians face when trying to piece together historical narratives about engineers. Unlike scientists and artists, he notes, engineers tend not to write memoirs, nor do they show up in archival collections.

"The signal strength tends to be swamped by artists or scientists," McCray says.

In "Big in Japan," in this issue, he writes about one engineering project that did leave a vast paper trail: building the Pepsi Pavilion for the 1970 World's Fair, in Osaka. That extensive record may have something to do with the fact that the team who built the pavilion included numerous artists and scientists as well as engineers. The pavilion was the work of Experiments in Art and Technology, or E.A.T., an influential group overseen by Bell Telephone Laboratories engineer Johan Wilhelm "Billy" Klüver.

Over the years, art critics and art historians have considered the pavilion, but not from the engineers' perspective. McCray, a history professor at the University of California, Santa Barbara, whose academic degrees are all in engineering, enjoyed diving into the technical details and learning how Klüver and the other engineers collaborated with their artistic partners.

The Pepsi Pavilion is one of a number of art-and-technology projects that McCray features in his new book, *Making Art Work: How Cold War Engineers and Artists Forged a New Creative Culture*, which will be published this fall by the MIT Press. (The photo above shows McCray with a kinetic sculpture by the pioneering rocketeer Frank Malina, who is also profiled in the book.)

Part of McCray's research for the chapter on the Pepsi Pavilion involved "plowing through two dozen giant boxes of paper" archived at the Getty Research Institute. But he also contacted people involved with the project to get access to materials that weren't in formal collections. For example, Julie Martin, Klüver's widow, shared with McCray an extensive set of personal records and photos.

"It's like you're a detective in a police procedural gathering evidence," McCray says. "After a while you end up with a new view." ■

## CONTRIBUTORS

### Edmon de Haro

De Haro is a graphic designer based near Barcelona. To illustrate "AI: Expect Evolution, Not Revolution" [p. 30], he sought to convey the fact that "right now, artificial intelligence may not be as useful or as powerful as people think," de Haro says. "It's a path that will be very slow and difficult." In one image he depicts a snail climbing an incline, while a second image shows the evolution of humans, with a stumble at the end.

### Jeffrey Funk

Funk retired from the National University of Singapore in 2016. He now consults in various areas of technology and business, including the economic effects of artificial intelligence, which he writes about in this issue [p. 30]. His interest in the impact of technology on productivity began in the 1980s with his doctoral work on the economics of industrial robots. "The rapid diffusion of robots has always been overhyped," says Funk. "And AI will likely be the same."

### Shaoshan Liu

Liu is the founder and CEO of PerceptIn, an autonomous vehicle startup in Fishers, Ind. In this issue, he and Jean-Luc Gaudiot, a professor at the University of California, Irvine, describe PerceptIn's efforts to develop low-speed autonomous passenger shuttles [p. 36]. "Riding the first time in PerceptIn's autonomous vehicle was magical," says Liu. Although it was traveling at low speed in an area with limited traffic, he says, "it was like the moon-landing moment."

### Pete Lewis

After earning a music degree, Lewis designed audio equipment for his band and discovered a passion for electrical engineering. This led him to a career at SparkFun Electronics, where he started as an assembly technician and now works as quality control manager. In this issue [p. 16], he discusses how to use a crypto chip in a DIY smart home and other applications. "It's fun to think that now, with the coprocessor, you can add a really strong layer of security to all your projects," says Lewis.

### Peter Palomaki

Palomaki is the owner and chief scientist of Palomaki Consulting, where he helps companies implement quantum-dot technology. A prime example is the quantum-dot image sensors that he and Sean Keuleyan, lead scientist at Voxtel, write about in "Snapshots by Quantum Dots" [p. 24]. Passionate about all things QD, Palomaki once made Christmas decorations using red and green quantum dots. The QDs contained cadmium, so he used them only in the lab.


# HERE COMES DRIVERLESS RIDE SHARING

## Cruise unveils the Origin, a fully autonomous SUV designed for app-controlled urban transportation

I recently drove from Silicon Valley to San Francisco. It started raining on the way and I hadn't thought to take an umbrella. No matter—I had the locations of two parking garages, just a block or so from my destination, preloaded into my navigation app. But both were full, and I found myself circling in stop-and-go traffic around crowded, wet, hilly, construction-heavy San Francisco, hunting for street parking or an open garage for nearly an hour. It was driving hell. ¶ So when I finally arrived at a launch event hosted by Cruise, I couldn't have been more receptive to the company's pitch for the Cruise Origin, a new vehicle that, Cruise executives say, intends to make it so I won't need to drive or park in a city ever again. ¶ The Cruise Origin is a six-passenger, autonomous, electric, SUV-size vehicle intended to disrupt not so much the car industry as urban transportation overall. Cruise (mostly owned by GM) does not plan to offer the Origin on the retail market. Instead, it will operate fleets of the vehicles as a ride-sharing service; screens inside are intended to give information about upcoming pickups and drop-offs. ¶ Uber, which launched the last big transportation disruption and has been preparing for the next by investing in its own autonomous vehicle research, might have some scrambling to do. ¶ Since the Origin won't be sold, the company isn't talking about pricing. ¶ However, Cruise CEO Dan Ammann did talk a lot about what the designers did to make this autonomous vehicle as inexpensive as possible to manufacture—production costs will be about half of those required to make today's all-electric SUVs, he said. The designers started with a new, all-electric platform, made all the sensor and computer systems modular for easy replacement and upgrading, and took out everything driver-related, including rearview mirrors, windshield wipers, and, of course, the steering wheel.

**CRUISE'S CTO,** Kyle Vogt, presents the Origin, a driverless electric shuttle.

Besides reducing costs, those omissions left room for a big passenger compartment. I do have one quibble with the design, though: in the display vehicle, passengers faced each other in two rows of seats with lots of room in between. While this arrangement might be great for Vegas party limos, those of us who are motion sensitive need to face the front and have good sight lines in the direction of travel. And, frankly, even if I weren't motion sensitive, I don't necessarily want to spend my travel time awkwardly avoiding eye contact with a stranger.

"It costs a lot less to make than you would expect, it will last a million miles, and you can share it," Ammann said. The company estimates that the average urban dweller who relies on Cruise Origin for transportation will cut about US $5,000 a year from personal transportation costs. "The key to making money is making a better user experience at a lower cost."

The vehicle is "fully engineered and on its way to production," Ammann said. Operating it as a driverless service still needs government approvals, however. —TEKLA S. PERRY

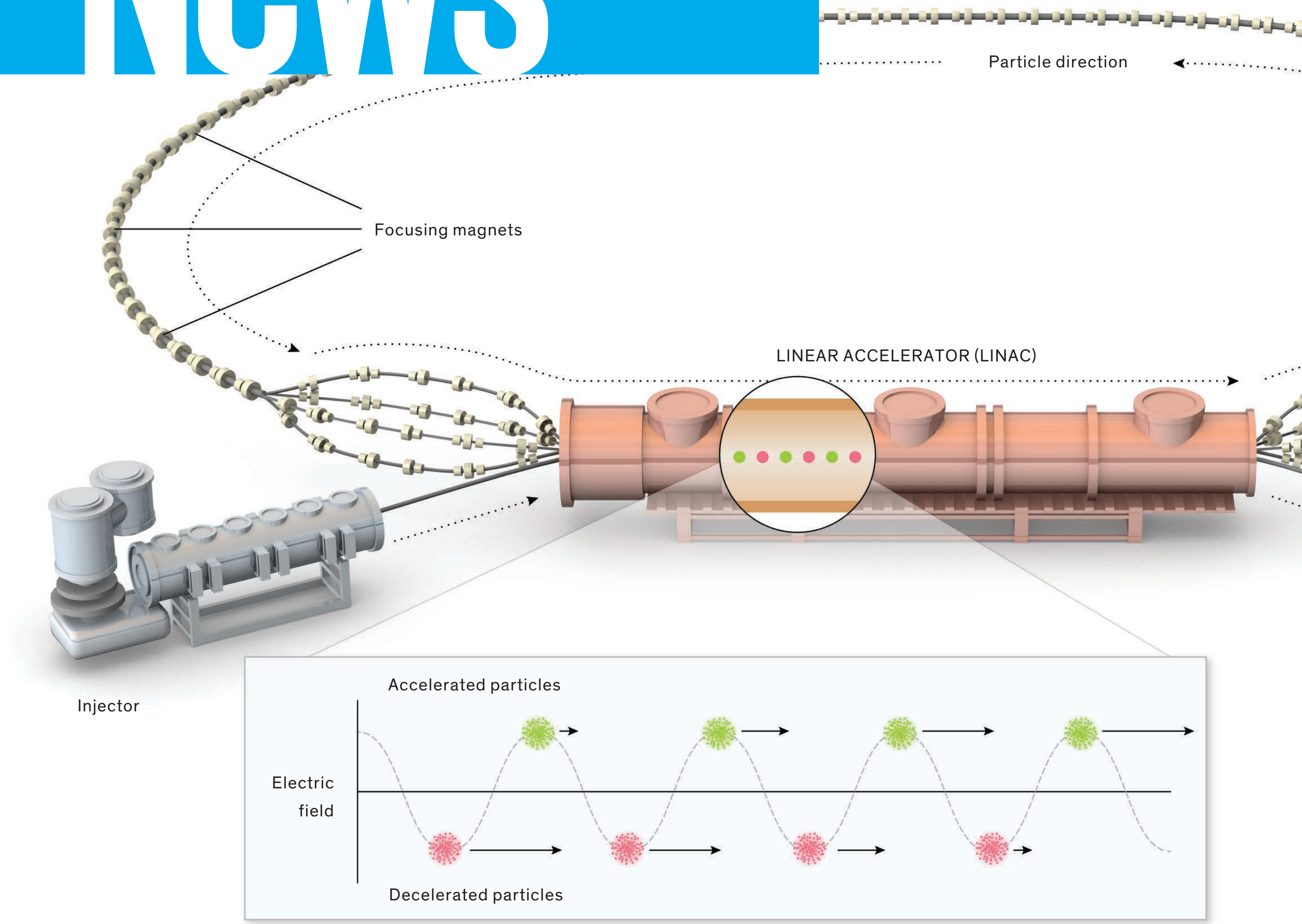
*A version of this article appears in our View From the Valley blog.*

**POST YOUR COMMENTS AT** spectrum.ieee.org/spectrallines-mar 2020

**CORRECTION:** In "Transgenic Salmon Hits U.S. Shelves" [January], we misstated the date by which U.S. food manufacturers must apply new labels to bioengineered foods. The correct date is January 2022.

DAVID PAUL MORRIS/BLOOMBERG/GETTY IMAGES

# News




# NEW PARTICLE-ACCELERATING TECH PASSES TEST

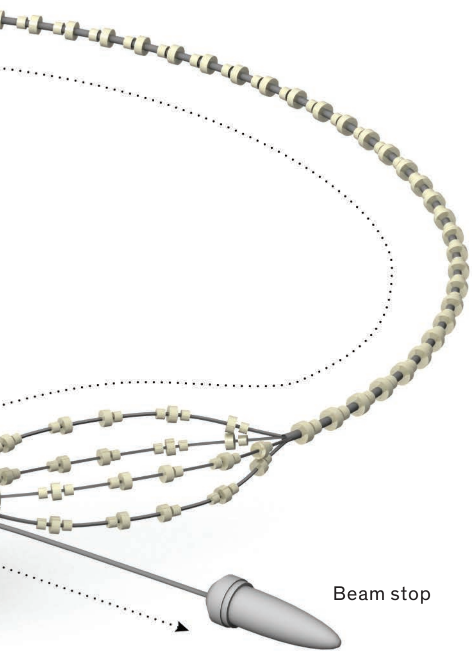
## Energy recovery linear accelerators could reduce costs and conserve power

**Accelerator physicists from Cornell** University and Brookhaven National Laboratory have facilitated an unprecedented energy handoff between electrons.

As particle colliders have gotten bigger and more expensive to build and operate, physicists have begun to look for nontraditional ways to accelerate particles. One potential solution is to use energy recovery linear accelerators, or ERLs. These new particle accelerators transfer energy from decelerating electrons to give fresh particles a boost–similar to the way speed skaters transfer energy by physi-



ILLUSTRATION BY Emily Cooper




**HOW ENERGY RECOVERY WORKS: Particles injected into the linac of an energy recovery linear accelerator are propelled forward by an electric field, then steered by magnets around a curve. Along the way, the particles emit synchrotron radiation. When they return to the linac, the particles are accelerated again. After several trips, they decelerate, transferring kinetic energy back to the electric field. This energy accelerates new particles. Decelerated particles are discarded into the beam stop.**

cally pushing their teammates forward to begin each new leg of a relay.

CBETA (short for Cornell-BNL ERL Test Accelerator) is a proof-of-concept experiment for such next-generation accelerating technology. Last December, researchers managed to achieve what's called eight-pass energy recovery for CBETA, a benchmark that shows the technology's potential for future colliders.

Conventional particle accelerators fall into one of two main classes: linear accelerators or storage rings. Linear accelerators, also known as linacs, are hollow metal chambers filled with strong electric fields. These fields flip on and off, and with the right timing, charged particles inside the chambers can be propelled forward or backward. The resulting particle beam is dense but has relatively few particles.

Storage rings circulate particles millions of times by bending their paths with magnets. Particles can be continually injected into a storage ring, which creates a beam with more particles. But the beam has lower density, becoming diluted as the particles circulate.

Georg Hoffstaetter, a physics professor at Cornell who leads CBETA, says ERLs combine the strengths of both. "We have two traditional accelerator technologies: linacs, which can provide low current but very dense beams, and rings, which can provide high current but less dense beams," he says. "An ERL merges these two technologies to get both advantages—to get high currents for very dense beams."

Trying to make two beams of particles collide results mostly in misses because the particles are incredibly small. Physicists love dense beams and high currents because both qualities provide more collisions and therefore more data.

The concept of an ERL has been around since 1965, when Cornell physicist Maury Tigner proposed it, but the technology has become attractive only in recent years, in part because of how complex the energy handoff is to execute.

In ERLs, particles are initially accelerated by a linear accelerator. Magnets then "loop" the particles back to the beginning so that they pass through the linear accelerator again. In CBETA, electrons make eight full passes. On the first four, the electrons gain energy. But after the fourth pass, they arrive out of sync, and the electric field, instead of pushing them forward, slows them down.

As with speed skaters, when these electrons slow down they lose their kinetic energy. But energy is conserved—it has to go somewhere. For skaters, the energy moves through a push to the next skater; for electrons, the energy moves through the electric field to the next accelerating electron. After an electron finishes its fourth deceleration, it's discarded.

Because they combine the advantages of both linacs and storage rings, ERLs present a tempting alternative to current collider tech. Besides CBETA, a few other ERLs have achieved full energy recovery, but not for eight passes. More passes give the electrons higher energy, but this also makes the particles more difficult to control.

"ERLs are notoriously hard to commission, and the fact that they've managed eight-pass recovery using permanent magnets is quite a feat," says Ryan Bodenstein, an accelerator physicist at the Belgian Nuclear Research Centre. "I'm really quite excited about this breakthrough."

The European, Japanese, and American particle-physics communities are deliberating what future accelerators to fund. CBETA's success may cause them to take another look at ERLs—which, thanks to their smaller size and power savings, reduce costs. Some future experiments, such as an electron-ion collider to be built at Brookhaven, will use ERLs.

ERLs still face challenges, though. There are questions about whether the handoff would go as smoothly in a real collider: Smashing beams of electrons with ions or other particles could throw off the timing of the sensitive energy handoff. Design complications could take years to smooth out.

"I think the ideas should be pursued and investigated further," says Bodenstein. "And even if it doesn't really work out in this case, I think it will provide some great insights." —DAN GARISTO

**POST YOUR COMMENTS AT**
spectrum.ieee.org/accelerator-mar2020

NEWS

JOURNAL WATCH

## The Boredom Detector

**A PROFESSOR FINISHES** a lecture and checks his computer. A software program shows that most students lost interest after 30 minutes. The professor makes a note to revise the lecture.

Scientists are working to make this fictional scenario a reality. In a paper recently published in *IEEE Transactions on Visualization and Computer Graphics*, researchers describe an artificial intelligence (AI) system that analyzes students' emotions based on video recordings of their facial expressions.

The system "provides teachers with a quick and convenient measure of the students' engagement level in a class," says Huamin Qu, a computer scientist at the Hong Kong University of Science and Technology.

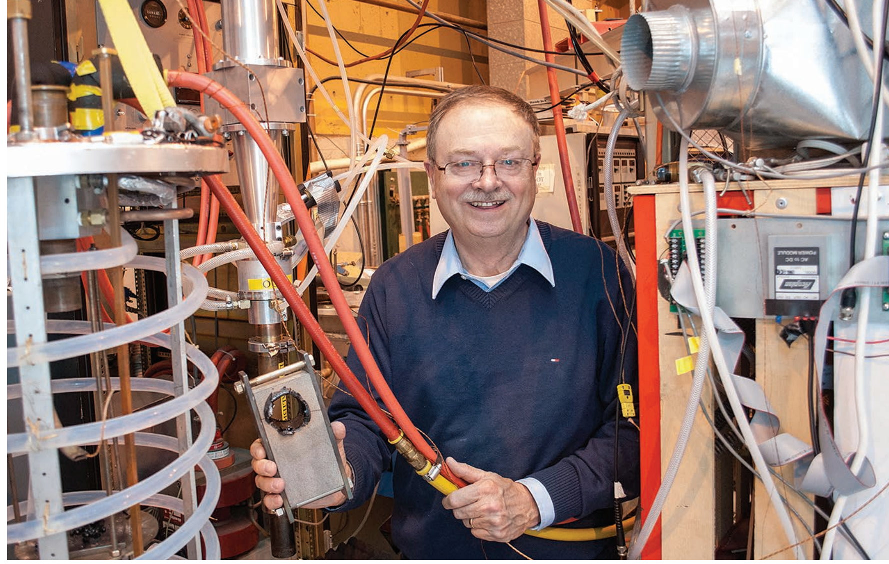
Qu and his colleagues tested their AI system in two classrooms consisting of toddlers in Japan and university students in Hong Kong. The visual analytics program did a good job of detecting happiness. But the model often incorrectly reported anger or sadness when students were actually just focused.

The "focus frown" and other confusing facial expressions are a challenge for just about everyone working in the field of emotion recognition, says Richard Tong, chief architect at Squirrel AI Learning. "We have had similar problems in our own experiments," he says.

Plus, putting video cameras in the classroom creates privacy issues. And if the cameras distract students or teachers, the plan could backfire.

Instead, Tong envisions using emotion recognition to develop AI tutors. These computer-based teachers will be trained to spot when a student is losing interest. The AI can then adjust its teaching strategy accordingly. —**EMILY WALTZ**

*An extended version of this article appears on our website in the Journal Watch section.*



PAUL RIVENBERG/MIT

# ALTAROCK MELTS ROCK FOR GEOTHERMAL WELLS

### Millimeter waves could help us dig deeper and faster than with traditional drills

**A vast supply of heat lies** beneath our feet. Yet today's drilling methods can barely push through dense rocks and high-pressure conditions to reach it. A new generation of "enhanced" drilling systems aims to obliterate those barriers and unlock unprecedented supplies of geothermal energy.

AltaRock Energy is leading an effort to melt and vaporize rocks with millimeter waves. Instead of grinding away with mechanical drills, scientists use a gyrotron—a specialized high-frequency microwave-beam generator—to open holes in slabs of hard rock. The goal is to penetrate rock at faster speeds, to greater depths, and at a lower cost than conventional drills do.

The Seattle-based company recently received a US $3.9 million grant from the U.S. Department of Energy's Advanced Research Projects Agency-Energy (ARPA-E). The three-year initiative will enable scientists to demonstrate the technology at increasingly larger scales, from burning through hand-size samples to room-size slabs. Project partners say they hope to start drilling in real-world test sites before the grant period ends in September 2022.

AltaRock estimates that just 0.1 percent of the planet's heat content could supply humanity's total energy needs for 2 million years. Earth's core, at a scorching 6,000 °C, radiates heat through layers of magma, continental crust, and sedimentary rock. At extreme depths, that heat is available in constant supply anywhere on the planet. But most geothermal projects don't reach deeper than 3 kilometers, owing to technical or financial restrictions. Many wells tap heat from geysers or hot springs close to the surface.

That's one reason why, despite its potential, geothermal energy accounts for only about 0.2 percent of global power capacity, according to the International Renewable Energy Association.

"Today we have an access problem," says Carlos Araque, CEO of Quaise, an

NEWS

**FEEL THE HEAT:** Paul Woskov of MIT holds water-cooling lines leading to a test chamber, and a sample of rock with a hole made by a beam from a gyrotron.

affiliate of AltaRock. "The promise is that, if we could drill 10 to 20 km deep, we'd basically have access to an infinite source of energy."

The ARPA-E initiative uses technology first developed by Paul Woskov, a senior research engineer at MIT's Plasma Science and Fusion Center. Since 2008, Woskov and his colleagues have used a 10-kilowatt gyrotron to produce millimeter waves at frequencies between 30 and 300 gigahertz. Elsewhere, millimeter waves are used for many purposes, including 5G wireless networks, airport security, and astronomy. While producing those waves requires only milliwatts of power, it takes several megawatts to drill through rocks.

To start, MIT researchers place a piece of rock in a test chamber, then blast it with high-powered, high-frequency beams. A metallic waveguide directs the beams to form holes. Compressed gas is injected to prevent plasma from breaking down and bursting into flames, which would hamper the process. In trials, millimeter waves have bored holes through granite, basalt, sandstone, and limestone.

The ARPA-E grant will allow the MIT team to develop their process using megawatt-size gyrotrons at Oak Ridge National Laboratory, in Tennessee. "We're trying to bring forward a disruption in technology to open up the way for deep geothermal energy," Araque says.

Other enhanced geothermal systems now under way use mechanical methods to extract energy from deeper wells and hotter sources. In Iceland, engineers are drilling 5 km deep into magma reservoirs, boring down between two tectonic plates. Demonstration projects in Australia, Japan, Mexico, and the U.S. West—including one by AltaRock—involve drilling artificial fractures into continental rocks. Engineers then inject water or liquid biomass into the fractures and pump it to the surface. When the liquid surpasses 374 °C and 22,100 kilopascals of pressure, it becomes a "supercritical" fluid, meaning it can transfer energy more efficiently and flow more easily than water from a typical well.

However, such efforts can trigger seismic activity, and projects in Switzerland and South Korea were shut down after earthquakes rattled surrounding cities. Such risks aren't expected for millimeter-wave drilling. Araque says that while beams could spill outside their boreholes, any damage would be confined deep below ground.

Maria Richards, coordinator at Southern Methodist University's Geothermal Laboratory, in Dallas, says that one advantage of using millimeter waves is that the drilling can occur almost anywhere—including alongside existing power plants. At shuttered coal facilities, deep geothermal wells could produce steam to drive the existing turbines.

The Texas laboratory previously explored using geothermal power to help natural-gas plants operate more efficiently. "In the end, it was too expensive. But if we could have drilled deeper and gotten higher temperatures, a project like ours would've been more profitable," Richards says. She notes that millimeter-wave beams could also reach high-pressure offshore oil and gas reservoirs that are too dangerous for mechanical drills to tap.
—MARIA GALLUCCI

**POST YOUR COMMENTS AT**
spectrum.ieee.org/geothermal-mar2020

# DATA PROJECT AIMS TO ORGANIZE SCIENTIFIC RECORDS

## Deep-time Digital Earth will link hundreds of bespoke databases

**Geoscience researchers are** excited by a new big-data effort to connect millions of hard-won scientific records in databases around the world. When complete, the network will be a virtual portal into the ancient history of the planet.

The project is called Deep-time Digital Earth, and one of its leaders, Nanjing-based paleontologist Fan Junxuan, says it unites hundreds of researchers—geochemists, geologists, mineralogists, paleontologists—in an ambitious plan to link potentially hundreds of databases.

The Chinese government has lined up US $75 million for a planned complex near Shanghai that will house dedicated programming teams and academics supporting the project, and a supercomputer for related research. More support will come from other institutions and companies, with Fan estimating total costs to create the network at about $90 million.

Right now, a handful of independent databases with more than a million records each serve the geosciences. But there are hundreds more out there holding data related to Earth's history. These smaller collections were built with assorted software and documentation formats. They're kept on local hard drives or institutional servers, some decades old, and converted from one format into another as time,

**FOSSIL RECORD:** Deep-time Digital Earth will make it easier for scientists to study fossils such as these. The project is led by paleontologist Fan Junxuan [right].

funding, and interest allow. The data might be in different languages and is often guided by informal or variably defined concepts. There is no standard for arranging the hundreds of tables or thousands of fields. This archipelago of information is potentially very useful but hard to access.

Fan saw an opportunity while building a database comprising the Chinese geological literature. Once it was complete, he and his colleagues were able to use parallel computing programs to examine data on 11,000 marine fossil species in 3,000 geological sections. The results dated patterns of paleobiodiversity—the appearance, flowering, and extinction of whole species—at a temporal resolution of 26,000 years. In geologic time, that's pretty accurate.

The Deep-time project planners want to build a decentralized system that would bring these large and small data sources together. The main technical challenge is not to aggregate petabytes of data on centralized servers but rather to script strings of code. These strings would work through a programming interface to link individual databases so that any user could extract information through that interface.

Harmonizing these data fields requires human beings to talk to one another. Fan and his colleagues hope to kick off those discussions in New Delhi, which this month is hosting a big gathering of geoscientists. A linked network could be a gold mine for researchers scouring geologic data for clues.

In a 19th-century building behind Berlin's Museum für Naturkunde, micropaleontology curator David Lazarus and paleobiologist postdoc Johan Renaudie run the group's Neptune database, which is likely to be linked with Deep-time Digital Earth as it develops. Neptune holds a wealth of data on core samples from the world's ocean floors. Lazarus started the database in the late 1980s, before the current SQL language standard was readily available—at that time it was mostly found only on mainframes. Renaudie explains that Neptune has been modified from its incarnation as a relational database using 4th Dimension for Mac, and has been carefully patched over the years.

There are many such patched-up archives in the field, and some researchers start, develop, and care for data centers that drift into oblivion when funding runs out. "We call them whale fall," Lazarus says, referring to dead whales that sink to the ocean floor.

Creating a database network could keep this information alive longer and distribute it further. It could lead to new kinds of queries, says Mike Benton, a vertebrate paleontologist in Bristol, England, making it possible to combine independent data sources with iterative algorithms that run through millions or billions of equations. Doing this can deliver more precise time resolutions, which hitherto has been really difficult. "If you want to analyze the dynamics of ancient geography and climate and its influence on life, you need a high-resolution geological timeline," Fan says. "Right now this analysis is not available." —MICHAEL DUMIAK

**POST YOUR COMMENTS AT**
spectrum.ieee.org/digitalearth-mar2020

NEWS

FAN JUNXUAN (3)

# POTASSIUM BATTERIES SHOW PROMISE

## Hurdles remain, but potassium could someday make sense for grid storage

**Renewables are poised to expand** by 50 percent in the next five years, according to the International Energy Agency. Much of that wind and solar power will need to be stored. But a growing electric-vehicle market might not leave enough lithium and cobalt for lithium-ion grid batteries.

Some battery researchers are taking a fresh look at lithium's long-ignored cousin, potassium, for grid storage. Potassium is abundant, inexpensive, and could in theory enable a higher-power battery. However, efforts have lagged behind research on lithium and sodium batteries.

But potassium could catch up quickly, says Shinichi Komaba, who leads potassium-ion battery research at the Tokyo University of Science: "Although potassium-battery development has just been going on for five years, I believe that it is already competitive with sodium-ion batteries and expect it to be comparable and superior to lithium-ion."

People have historically shied away from potassium because the metal is highly reactive and dangerous to handle. What's more, finding electrode materials to hold the much heftier potassium ions is difficult.

Yet a flurry of reports in the past five years detail promising candidates for the cathode. Among the leaders are iron-based compounds with a crystalline structure similar to Prussian blue particles, which have wide open spaces for potassium ions to fill. A group from the University of Texas at Austin led by John Goodenough, coinventor of the lithium-ion battery and a winner of the 2019 Nobel Prize in Chemistry, has reported Prussian blue cathodes with an exceptionally high energy density of 510 watt-hours per kilogram, comparable to that of today's lithium batteries.

But Prussian blue isn't perfect. "The problem is, we don't know how water content in the material affects energy density," says Haegyeom Kim, a materials scientist at Lawrence Berkeley National Laboratory. "Another issue is that it's difficult to control its chemical composition."

Kim is placing bets on polyanionic compounds, which are made by combining potassium with any number of elements plucked from the periodic table. Potassium vanadium fluorophosphate seems to hold special promise.

Kim and his colleagues have developed a cathode with the compounds that has an energy density of 450 Wh/kg.

Other researchers are looking at organic compounds for cathodes. These cost less than inorganic compounds, and their chemical bonds can stretch to take up potassium ions more easily.

While Goodenough is giving potassium a chance, his fellow lithium-battery inventor and Nobel Prize winner M. Stanley Whittingham, professor of chemistry at Binghamton University, in New York, isn't sold. "It's a scientific curiosity," he says. "There's no startup looking at potassium batteries."

Potassium, says Whittingham, is not a practical technology because of its heft and volatility. Potassium also melts at a lower temperature than lithium or sodium, which can trigger reactions that lead to thermal runaway.

Those are valid concerns, says Vilas Pol, a professor of chemical engineering at Purdue University, in West Lafayette, Ind. But he points out that in a battery, potassium ions shuttle back and forth, not reactive potassium metal. Special binders on the electrode can tame the heat-producing reactions.

Developing the right electrolyte will be key to battery life and safety, says Komaba, of the Tokyo University of Science. Conventional electrolytes contain flammable solvents that, when combined with potassium's reactivity, could be dangerous. Selecting the right solvents, potassium salts, salt concentration, and additives can prevent fires.

Komaba's group has made electrolytes using potassium-fluoride salts, superconcentrated electrolytes that have fewer solvents than traditional mixes, and ionic liquid electrolytes that don't use solvents. In January, materials scientist Zaiping Guo and her team from the University of Wollongong, Australia, reported a nonflammable electrolyte for potassium batteries. They added a flame retardant to the solvent.

Potassium enthusiasts point out that the technology is still at an early stage. It's never going to match the high energy density of lithium, or be suitable for electric cars. Yet for immense grid batteries, cheap potassium might have an upper hand. "Potassium-ion [batteries] could have worked earlier, but there was no need for [them]," says Pol. "Lithium isn't enough now."

In the end, the sum will have to be as good as its parts. Most research has focused on the materials that go into the electrodes and the electrolyte. Put it all together in a battery cell and the energy density drops after just 100 charging cycles or so; practical batteries will need to withstand several hundred.

"It will take time to figure out the exact combination of electrolyte, cathode, and anode," Pol says. "It might take another 15 years from now to get to the market."

—PRACHI PATEL

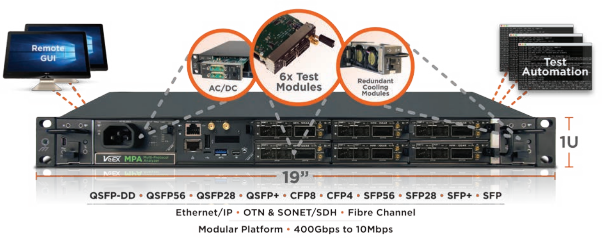
**POST YOUR COMMENTS AT**
spectrum.ieee.org/potassium-mar2020

# 400G: Testing the Future of Communications


As PAM4-based 400GE QSFP-DD and OSFP transceivers go into full commercial deployment, testing and verification needs change and move from the pure R&D labs, SVT, manufacturing, FAEs supporting demonstrations and field evaluations to field deployment. Not all 400G test and measurement applications are the same. Whether it is a portable do-it-all handheld for the field, high mobility and availability in large datacenters, high port density rack-mount environments, or automation, VeEX® solutions offer the right products and features to address applications from R&D to manufacturing, from Core to Field. Our comprehensive test and measurement portfolio of products and applications fulfill verification needs, from layer 1 to 4, from 10M to 400G, from lab to field.

## Test Solutions for High Speed Interfaces & Transport Networks

- **Applications** - Network equipment, systems and IC development; signal integrity verification; transceiver validation; design verification and system testing (SVT); production and manufacturing test; network verification and service delivery
- **Portability** - First 400GE handheld test set, with Native PAM4 QSFP-DD support for best-in-class signal integrity. Enables lab-to-field transition, transceiver verification, benchmarking and link commissioning and troubleshooting
- **Scalability** - Rack-mount applications for high-port count, manufacturing and automation applications
- **Flexibility** - 400GE, 200GE, 100GE, 50GE, 40GE, 25GE, 10GE, FlexE, OTUCn, OTU4, OTU3, and more
- **Future proof** - Modular test platforms supporting current and future interfaces, rates and technologies
- **Pluggable Transceiver Interfaces** - OSFP, QSFP-DD, QSFP56, QSFP28, QSFP+, CFP8, CFP4, CFP2, SFP56, SFP28, SFP+, SFP
- Also available CPRI 10, 25G eCPRI, and 32G Fibre Channel, as well as lower-rate Ethernet, OTN, and legacy SDH/SONET PDH/DSn
- Advanced optical transceiver testing features


VeEX
The Verification Experts

Website: www.veexinc.com
Email: info@veexinc.com
Tel: +1.510.651.0500
Fax: +1.510.651.0505

**400G** *test solutions*
Visit us at OFC® - Booth #2614 and #4943

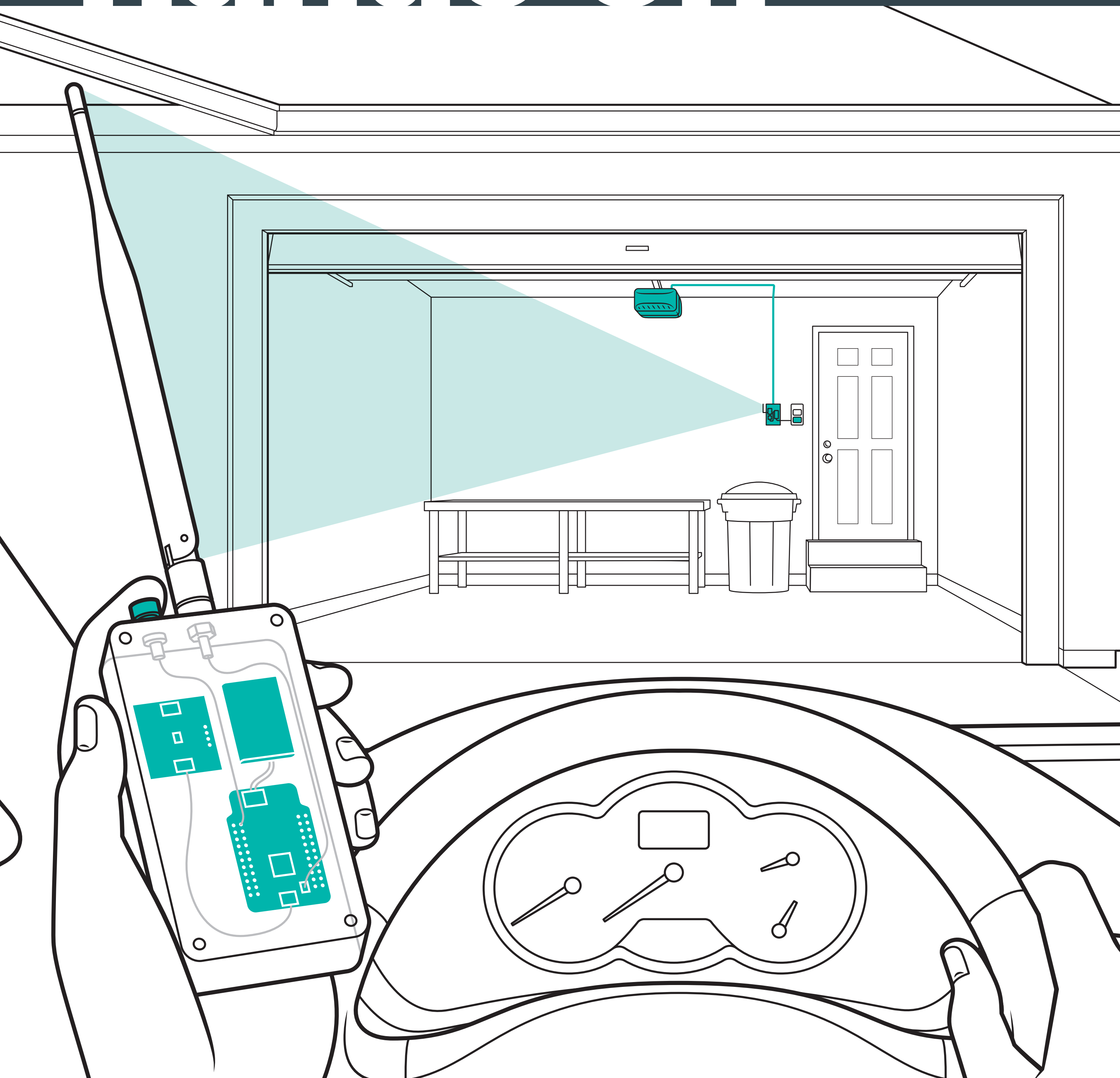


PHOTOGRAPH BY Rolls Royce/AP

# FLYING FAST, FAR, AND FUELLESS

**THIS PLANE,** scheduled to take its maiden flight this spring, represents the leading edge of the push to make electric aircraft fly as fast and as far as their fossil-fueled counterparts. The team, put together by luxury automobile maker Rolls Royce, has taken numerous design cues from Formula E racing planes. The one-seater's three lithium-ion batteries pack in enough energy (165 watt-hours per kilogram) to fly from London to Paris—about 320 kilometers (200 miles)—on a single charge. What's more, the plane will reach speeds approaching 480 kilometers per hour (300 miles per hour), which would shatter the 335 km/h (210 mph) record for electric planes.

# Hands On



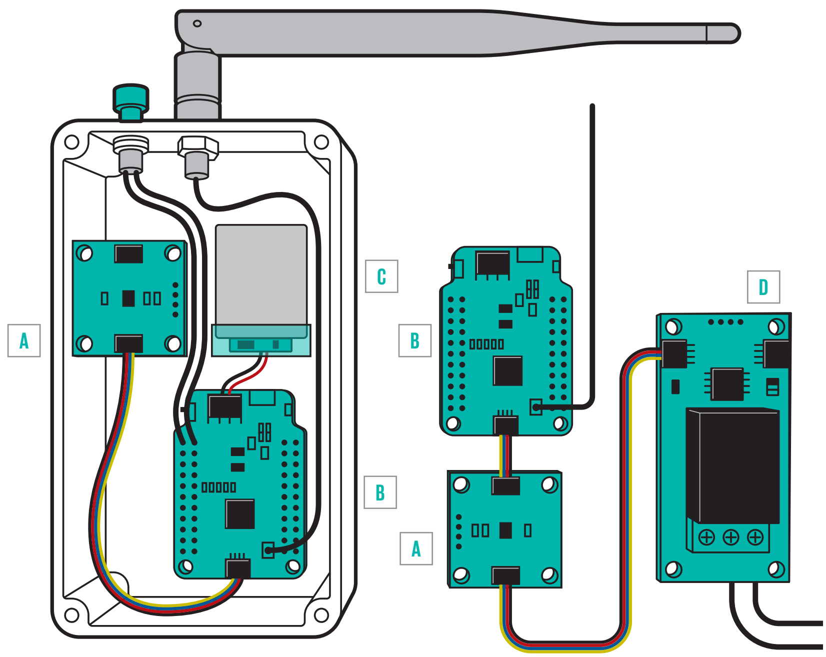
HANDS ON BY PETE LEWIS



ILLUSTRATIONS BY James Provost

# MAKE A HACK-PROOF GARAGE DOOR OPENER

## A NEW BREAKOUT BOARD OFFERS CRYPTOGRAPHIC SECURITY



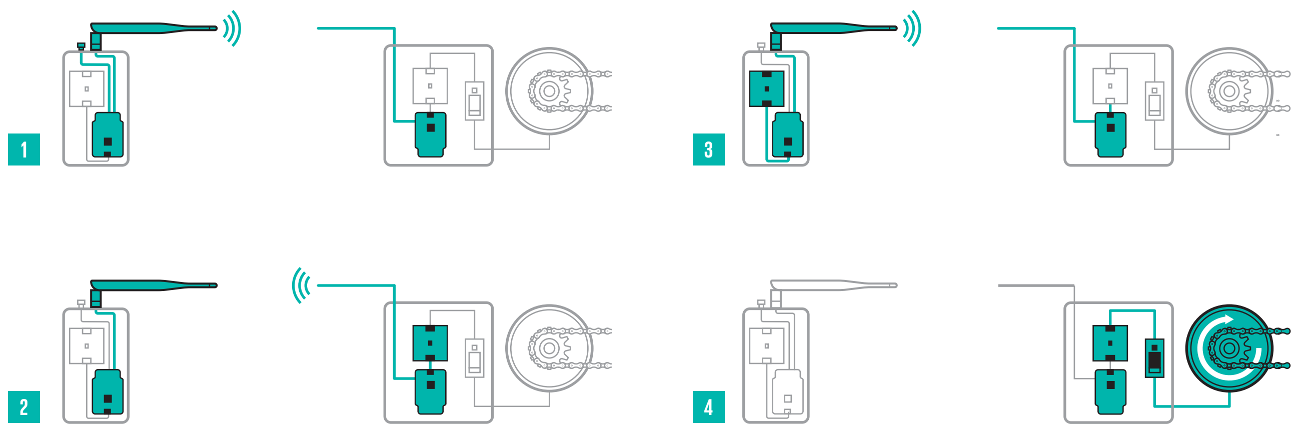

**CRYPTIC COPROCESSOR:** The ATECC508A coprocessor board [A] is connected to the Pro RF [B] in the remote [left], powered by a lithium polymer battery [C]. In the base station, the coprocessor and Pro RF use the I2C bus to control a relay [D], which activates the garage door mechanism [not shown].

**I'VE BEEN MAKING ELEC-**tronics projects for 15 years, but strong security was something I always considered out of my reach. Consequently, a fear of getting hacked limited the types of projects I would pursue, especially Internet-connected devices. But in May of 2019, I was handed the job of designing a cryptographic product for my employer, SparkFun. Among other things, SparkFun designs and sells breakout boards that allow makers to easily incorporate the capabilities offered by various integrated circuits into their designs. Now SparkFun wanted a board that would provide an easy on-ramp into the world of hardware-based cryptography.

It had to be user-friendly and Arduino compatible, which meant sifting through the specs of a lot of cryptographic hardware. What functions should our board offer, and how should it implement them? Ultimately, I chose to focus on ECC (elliptical curve cryptography) digital signatures. I'll get into a quick explanation of what ECC is in a moment, but the appeal of digital signatures is that they have a great real-world equivalent—handwritten signatures—which makes them a good introduction to cryptography. And signatures are very useful in the world of embedded systems, especially for those communicating over an insecure channel, like a radio link.

I had an immediate test application: As I started my crypto research, I remembered that my garage door remote control had stopped working years ago. I had wanted to replace the system with something of my own design, but I was never confident I could make something secure. But now my research had an extra impetus.

Venturing into the world of cryptography was pretty daunting, but with enough reading I found my way to a few datasheets of chips that use ECC-based crypto. ECC is similar to the RSA encryption algorithm often used on




**OPEN SESAME:** When the power button is pressed, three "$" characters are transmitted to the base station [1], which sends back a randomly generated token [2]. The token is signed with a private key and the signature is sent back to the base station [3]. If the signature can be verified against the token and public key, the door motor is activated [4].

the Internet—both use what's called a trapdoor mathematical function, which is easy to do but very hard to reverse. In RSA's case, the trapdoor function is the multiplication of two large prime numbers. If you have just the product of the numbers, it's very hard to factorize that back to its constituent primes, but if you know one prime and the product, it's trivial to do division and recover the other prime. With a trapdoor function in hand, you can create a private key and a public key. Anything encrypted with the public key can be decrypted only with the private key, and vice versa. In ECC's case, the trapdoor function is a hairy bit of math that exploits properties of points along an elliptic curve described by a formula of the form $y^2 = x^3 + ax + b$. If you're willing to take on the math, ECC lets you use shorter keys than RSA does, so it's better for embedded devices with limited power and bandwidth budgets.

After quite some searching, and following the advice of Josh Datko at Cryptronix, I came to the ATECC508A chip. It can do ECC signature creation and verification and talks I2C, the two-wire communications bus protocol that is well suited for Arduino compatibility. Time to order some samples!

The printed-circuit-board layout was fairly straightforward, and I had prototypes in no time. I plugged one in to my nearest Arduino, and it popped up on the correct I2C address. The hardware was verified. Now it was time for the difficult stuff: software!

The biggest hurdle was configuration. The ATECC508A has 126 configuration registers and there are many dependencies. If you attempt to change one thing, you often break another. Plus, in order to ensure the system is secure, once a configuration is chosen, it gets irreversibly locked: You only get one chance with these security ICs, so if you mess it up, then your IC is useless. Working very slowly and carefully, I nevertheless bricked several ICs (proud to say I never hit double digits). But I eventually found a suitable configuration that allowed for ECC signatures and verification. Whew! Finally I could begin writing examples for an Arduino library, demonstrating things like how to sign messages.

Now that the cryptographic coprocessor was completed, it was time to focus on fixing my garage door remote. The next big step was to add wireless communication. I opted to use a pair of SparkFun Pro RFs. They were nice to work with because they use an SAMD21 microcontroller with an I2C buffer large enough to handle the communications needs of the crypto coprocessor, and they have an onboard LoRa wireless transceiver, the RFM95. I initialized a crypto coprocessor, which creates a permanent private key—locked inside the coprocessor—and a public key which I could download via the I2C connection. (Step-by-step construction instructions and a bill of materials are available from the SparkFun site.)

I housed my remote in a sturdy aluminum case with a duck antenna and a single push button. Internally, it consists of my initialized crypto coprocessor board, a Pro RF, and a rechargeable lithium polymer battery. The normally open push button is wired between the battery and the Pro RF, so the board is off most of the time. Pressing the button for three seconds gives the board enough time to start up and complete the entire sequence to open the garage.

The sequence plays out like this: After boot up, the remote sends the string "$$$" to the base station in the garage (consisting of the other Pro RF and another ATECC508A crypto board with a copy of my remote's public key). The base station creates a token of random data using its ATECC508A and broadcasts it. The remote receives this token and creates a signature by combining the token with its private key, and transmits the signature. The base verifies the signature using the remote's public key. The security comes from the fact that the only place in the world that contains the unique private key necessary to make a valid signature is inside the remote's coprocessor. If all is good (within a strict time window), then the base opens the garage.

Next up, I plan to venture into areas that I was previously uncomfortable with. Now with this coprocessor in my bag of tricks, and good security in my hands, I'm ready to take on even the most concerning of IoT devices: my front door lock. ■

**POST YOUR COMMENTS AT**
spectrum.ieee.org/crytpo-mar2020

# Careers

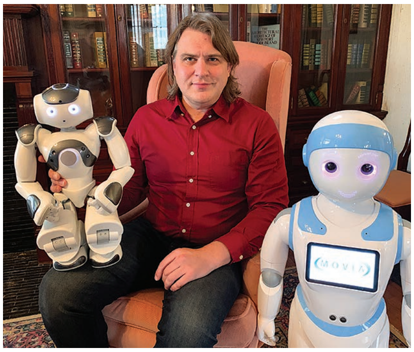
## TURNING ROBOTS INTO TEACHER'S AIDES

### CHRISTIAN WANAMAKER'S ROBOTICS SOFTWARE HELPS STUDENTS ON THE AUTISM SPECTRUM

**WHEN THE SMILEY-FACED RO**bot tells two boys to pick out the drawing of an ear from three choices, one of the boys, about 5, touches his nose. "No. Ear," his teacher says, a note of frustration in her voice. The child picks up the drawing of an ear and hands it to the other boy, who shows it to the robot. "Yes, that is the ear," the ever-patient robot says. "Good job." The boys smile as the teacher pats the first boy in congratulations.

The robot is powered by technology created by Movia Robotics, founded by Timothy Gifford in 2010 and headquartered in Bristol, Conn. Unlike other companies that have made robots intended to work with children with autism spectrum disorder (ASD), such as Beatbots, Movia focuses on building and integrating software that can work with a number of humanoid robots, including the Nao. Movia has robots in three school districts in Connecticut. Through a U.S. Department of Defense contract, its robots are being added to 60 schools for the children of military personnel worldwide.

It's Gifford's former computer science graduate student, Christian Wanamaker, who programs the robots. Before graduate school at the University of Connecticut, Wanamaker used his computer science degree to program commercial kitchen fryers. He enjoys a crispy fry as much as anyone, but his work coding for robot-assisted therapy is much more challenging, interesting, and rewarding, he says.




**FRIENDLY FACES:** Christian Wanamaker [middle] programs humanoid robots such as Nao [left] and iPal [right], so teachers can easily adapt them to the needs of students.

"I start with a robot that won't do anything without a programmer and end up with one that allows teachers to run therapies for children," he says. "That's very gratifying."

After graduating, he stayed on at the university as a research assistant. One of his first projects involved working as the lead developer writing code for an interactive media wall at Boston Children's Hospital as a way to give joy and control to sick kids. The multidiscipline team built a series of kid-friendly scenes designed to track movement and react. One scene on the three-story video screen displays grass swaying as someone passes by.

"That was pretty amazing, especially the response of the kids," he says.

Meanwhile, Gifford learned from his wife, a primary-school teacher, that the number of children in the classroom with ASD was growing, but staffing resources were limited. He collaborated with Wanamaker to program robots to work with ASD children in school in a nonthreatening way.

Gifford, a UConn researcher, talks with educators and clinicians about their students' needs and writes the software architecture to support the array of skills being taught. Gifford conveys this to Wanamaker, and it's up to him to break down the requirements into programming steps so that teachers can individualize the commands to suit each student's needs.

Wanamaker writes software using languages such as Python, Java, C#, and C++ so the robot can speak and move to direct the child as well as respond when the child reacts. In the early days, a team member controlled the robot; today, Movia has to ensure the robot can be controlled by a classroom teacher or other nontechnical person.

A coder working with robots needs curiosity, patience, and tenacity, Wanamaker says. Movia is constantly working with new robots that require him to use different languages and operating systems. Programming the robots to interact with individual people is not a straight line. "It's a bit like herding cats," he says.

"The thing I find rewarding about coding: You're literally creating something out of nothing," he says. "You're kind of like a wizard."

**—THERESA SULLIVAN BARGER**

*A version of this article appears in our Tech Talk blog.*

**POST YOUR COMMENTS AT** spectrum.ieee.org/asdrobots-Mar2020

MOVIA ROBOTICS

DEPARTMENTS

# CrossTalk

PM IMAGES/GETTY IMAGES

## CONCRETE FACTS

**THE ANCIENT** Romans were the first to mix sand and gravel with water and a bonding agent to make concrete. Although they called it *opus cementitium,* the bonding agent differed from that used in modern cement: It was a mixture of gypsum, quicklime, and pozzolana, a volcanic sand from Puteoli, near Mount Vesuvius, that made an outstanding material fit for massive vaults. Rome's Pantheon, completed in 126 C.E., still spans a greater distance than any other structure made of nonreinforced concrete.

The modern cement industry began in 1824, when Joseph Aspdin, of England, patented his firing of limestone and clay at high temperatures. Lime, silica, and alumina are the dominant constituents of modern cement; adding water, sand and gravel produces a slurry that hardens into concrete as it cures. The typical ratios are 7 to 15 percent cement, 14 to 21 percent water, and 60 to 75 percent sand and gravel.

Concrete is remarkably strong under compression. Today's formulations can resist a crushing pressure of more than 100 megapascals (14,500 pounds per square inch)—about the weight of an African bull elephant balanced on a coin. However, a pulling force of just 2 to 5 MPa can tear concrete apart; human skin is far stronger in this respect.

This tensile weakness can be offset by reinforcement. This technique was first used in iron-reinforced troughs for plants built by Joseph Monier, a French gardener, during the 1860s. Before the end of the 19th century, steel reinforcement was common in construction. In 1903 the Ingalls Building, in Cincinnati, became the world's first reinforced-concrete skyscraper. Eventually engineers began pouring concrete into forms containing steel wires or bars that were tensioned just before or after the concrete was cast. Such pre- or poststressing further enhances the material's tensile strength.

Today concrete is everywhere. It can be seen in the Burj Khalifa Tower in Dubai, the world's tallest building, and in the sail-like Sydney Opera House, perhaps

NUMBERS DON'T LIE BY VACLAV SMIL

the most visually striking application. Reinforced concrete has made it possible to build massive hydroelectric dams, long bridges, and gigantic offshore drilling platforms, as well as to pave roads, freeways, parking lots, and airport runways.

From 1900 to 1928, the U.S. consumption of cement (recall that cement makes up no more than 15 percent of concrete) rose tenfold, to 30 million metric tons. The postwar economic expansion, including the construction of the Interstate Highway System, raised consumption to a peak of about 128 million tons by 2005; recent rates are around 100 million tons a year. China became the world's largest producer in 1985, and its output of cement—above 2.3 billion metric tons in 2018—now accounts for nearly 60 percent of the global total. In 2017 and 2018 China made slightly more cement (about 4.7 billion tons) than the United States had made throughout the entire 20th century.

But concrete does not last forever, the Pantheon's extraordinary longevity constituting a rare exception. Concrete deteriorates in all climates in a process that is accelerated by acid deposition, vibration, structural overloading, and salt-induced corrosion of the reinforcing steel. As a result, the concretization of the world has produced tens of billions of tons of material that will soon have to be replaced, destroyed, or simply abandoned.

The environmental impact of concrete is another worry. The industry burns low-quality coal and petroleum coke, producing roughly a ton of carbon dioxide per ton of cement, which works out to about 5 percent of global carbon emissions from fossil fuels. This carbon footprint can be reduced by recycling concrete, by using blast-furnace slag and fly ash captured in power plants, or by adopting one of the several new low-carbon or no-carbon processes. But these improvements would make only a small dent in a business whose global output now surpasses 4 billion metric tons. ■

**POST YOUR COMMENTS AT**
spectrum.ieee.org/concrete-mar2020

THE NUMBERS GIVEN ARE THE WEIGHTS OF THE CONCRETE (NOT CEMENT) USED IN THE STRUCTURES' CONSTRUCTION.

1. THREE GORGES DAM, CHINA
65.5 MILLION METRIC TONS

2. GRAND COULEE DAM, WASHINGTON
21.7 MILLION METRIC TONS

3. PANAMA CANAL
6.8 MILLION METRIC TONS

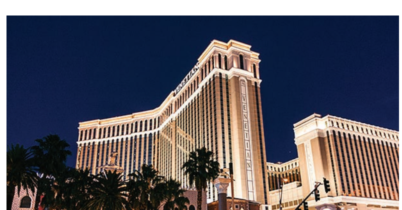
4. HOOVER DAM, ARIZONA AND NEVADA
6.0 MILLION METRIC TONS

5. KING FAHD CAUSEWAY, SAUDI ARABIA AND BAHRAIN, 0.84 MILLION METRIC TONS

6. THE PENTAGON, WASHINGTON, D.C.
0.80 MILLION METRIC TONS

7. PETRONAS TWIN TOWERS, MALAYSIA
0.39 MILLION METRIC TONS

8. BURJ KHALIFA TOWER, UNITED ARAB EMIRATES, 0.11 MILLION METRIC TONS

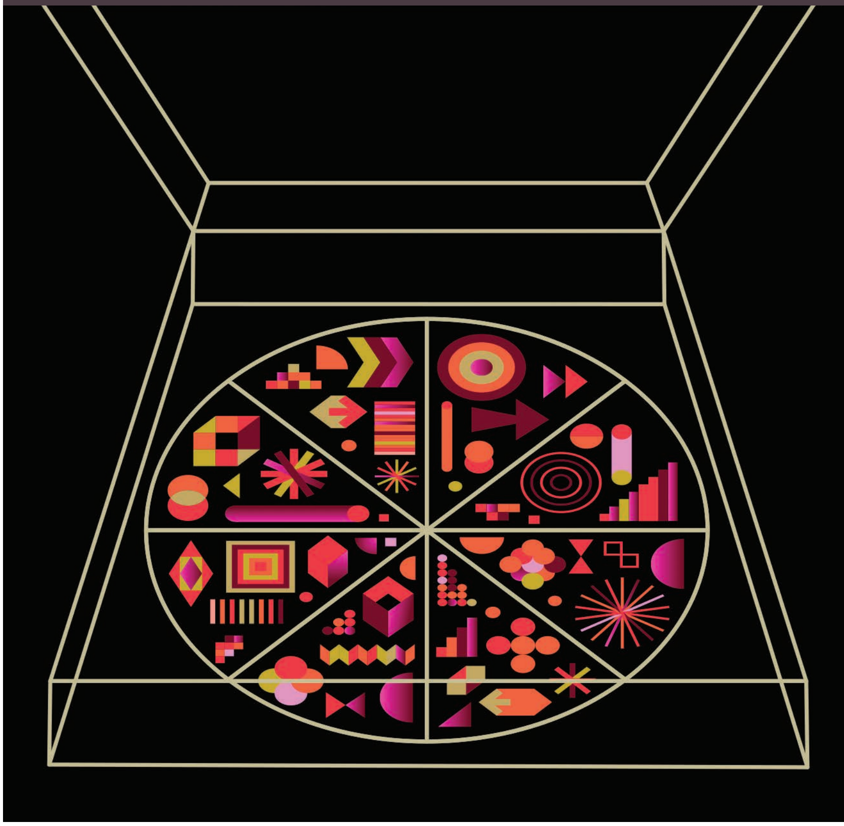


9. THE VENETIAN HOTEL, LAS VEGAS
0.039 MILLION METRIC TONS

10. WILSHIRE GRAND HOTEL, LOS ANGELES
0.037 MILLION METRIC TONS

DATA SOURCES: ARCHITIZER.COM, U.S. BUREAU OF RECLAMATION, PANAMA CANAL MUSEUM/UNIVERSITY OF FLORIDA; PHOTOS: ISTOCKPHOTO (10)

CROSSTALK



# WHAT 5G HYPE GETS WRONG

**THE HOTTEST FEATURE** of 5G isn't discussed very much. When people talk about 5G they tend to discuss the gigabit speeds or the lower latencies. But it's network slicing, the ability to partition off segments of the 5G network with specific latency, bandwidth, and quality-of-service guarantees, that could change the underlying economics of cellular service. Network slicing could lead to new companies that provide connectivity and help offset the capital costs associated with deploying 5G networks.

How? Instead of selling data on a per-gigabyte basis, these companies could sell wireless connectivity with specific parameters. A manufacturing facility, for example, may prioritize low latency so that its robots operate as precisely as possible. A hospital may want not only low latency but also dedicated bandwidth for telemedicine, to ensure that signals aren't lost at an inopportune moment.

Today, if a hospital or factory wants a dedicated wireless network with specific requirements, a telco has to custom-engineer it. But with network slicing, the telco can instead use software to allocate slices without human involvement. This would reduce the operating costs of a 5G network. That ease and flexibility, combined with the ability to price the network for different capabilities, will be what helps carriers justify the capital costs of deploying 5G, says Paul Challoner, the vice president of network product solutions for Ericsson North America.

Challoner envisions that soon customers will be able to go to a telco's website and define what they want, get the pricing for it, and then use the network slice for however long they need. He sees 2020 as being the year that equipment companies like Ericsson "race to the slice," trying to show wireless carriers what they can do.

Mobile-tech consultant Chetan Sharma thinks network slicing will likely take a year or two longer to hit the mainstream. But he also sees it as a catalyst for new companies that will enter the market to resell connectivity for dedicated use cases. For example, a company like Twilio or Particle, which already resell network connectivity to clients, could bring together slices from different carriers to offer a global service with specific characteristics. A company like BMW could then use that service when it wants to roll out a software update at a specific time to all of its vehicles—and to ensure that the update made it through.

Or maybe Amazon or Microsoft Azure could offer an industrial IoT product to factories that have specific latency requirements, by bundling together wireless connectivity from multiple carriers. A few years back, the telecom industry was debating whether carriers were becoming a dumb pipe. Sharma thinks the ability to customize speed, latency, and quality of service means 5G will put an end to that particular debate.

That said, carriers charging customers based on the capabilities they need does mean that some people will bring up concerns around network neutrality and how to ensure that customers aren't charged an arm and a leg for a decent best-effort service.

"It's uncharted territory," says Sharma. "When the FCC was looking at [network neutrality] they didn't consider network slicing as part of the equation. So my view is that they will have to update what operators are allowed to do with network slicing. We'll need more clarity on the ruling." ■

**POST YOUR COMMENTS AT**
spectrum.ieee.org/5ghype-mar2020



ILLUSTRATION BY Greg Mably

# THE COMPLACENCY TRAP

**AS I SWAM AROUND** the pool of blood, I said to myself, "There's a lesson here."

It hadn't felt that way when my lecture agents first invited me out for a "talk." Critiquing my skills as a public performer, they put it bluntly: "We're getting top dollar for you. And we think you can do better."

Better?

"You hide behind your iPad. Put it aside. Engage. Connect."

That advice landed like a ton of bricks, compacting my ego just enough for me to hear the truth in their words. I felt ashamed, embarrassed, and suddenly unqualified, tumbling from the top of the world to suffering impostor syndrome in less than a minute. I grew quiet, and my agents grew worried. I'd never truly learned how to take criticism without first feeling wounded. This time, with the breath knocked out of me, I chose to ignore the sting, setting my eyes on an opportunity: to be something more.

All of us walk a path throughout our lives. With a bit of luck, it leads to a comfortable destination, where we can make ourselves at home. Yet we need the occasional sleep on a bed of nails to remind us that we could benefit from some exercise. Movement keeps us trim, sharp, and healthy. Though we need rest, it should never be our goal.

Instead, take advantage of opportunities to walk with others, connecting and sharing and learning and teaching. For me, that means keeping pace with an enormous network of active and talented individuals from whom I can learn.

The day following that momentous meeting, a friend who also does public appearances recommended a class in improvisation. "It helped," she said. Before hearing those words, I'd never thought of needing theater skills for my craft—but of course I do. I enrolled in an improvisation workshop that evening. The following weekend I found myself swimming across that imaginary pool of blood, riffing off an idea offered by an improv partner.

Feeling now as though I've been jolted out of a lazy sleep, I hunger for more—for new skills, challenges, and opportunities. How can I be a better storyteller? Should I learn mime so that my body can tell the story? Voice-over skills? How to smile and speak to the camera? It feels like the first day of school, and I love it. So much to learn, so much to become. The best part: It never ends.

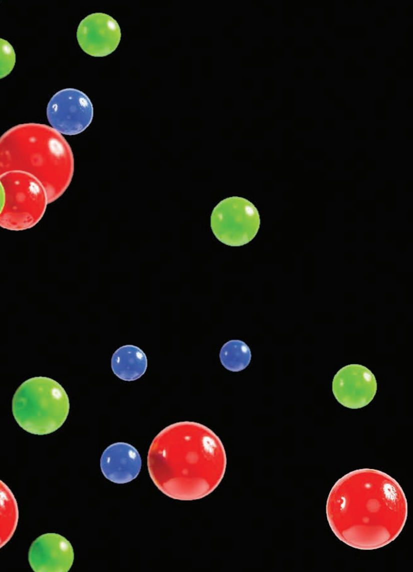
Though there could be more method to our growth. How often do we take the opportunity to reflect on what we can do well, then imagine what we want to be able to do? Can we write it out, naming it with, "This is where I excel, and here I fall short"? Putting ourselves in a place where we recognize our incompleteness may be uncomfortable, but it leaves us better able to imagine ourselves headed outward on a trajectory, making course corrections, toward an evolving destination. On my trajectory, that means acquiring theater skills. On yours, it could mean mastering millimeter-wave antenna design, or lidar, or memristors, or...

No one knows where we'll be in a year or a decade, but we have the power to decide for ourselves whether we'll be standing still or moving forward. With so many opportunities to connect with and learn from friends and colleagues, we need never remain in place. And if we remember to offer what we ourselves have learned, others will walk alongside, keeping stride, learning from us. ■

**POST YOUR COMMENTS AT**
spectrum.ieee.org/complacency-mar2020

ILLUSTRATION BY Dan Page


E 3.5-5.6/18-55 OSS
0.25m/0.82ft
Ø49



# Snapshots by Quantum Dots

## LOOK OUT, CMOS IMAGE SENSORS. THERE'S A NEW KID IN CAMERA-TOWN

**BY PETER PALOMAKI & SEAN KEULEYAN**

In the early 2000s, the commercialization of CMOS image sensors led to smaller and smaller—and cheaper and cheaper—digital cameras. Now the thinnest of mobile phones contains at least two camera modules, and all except the most dedicated photographers have stopped carrying a separate camera, concluding that the camera sensors in their phones take pictures that are good enough.

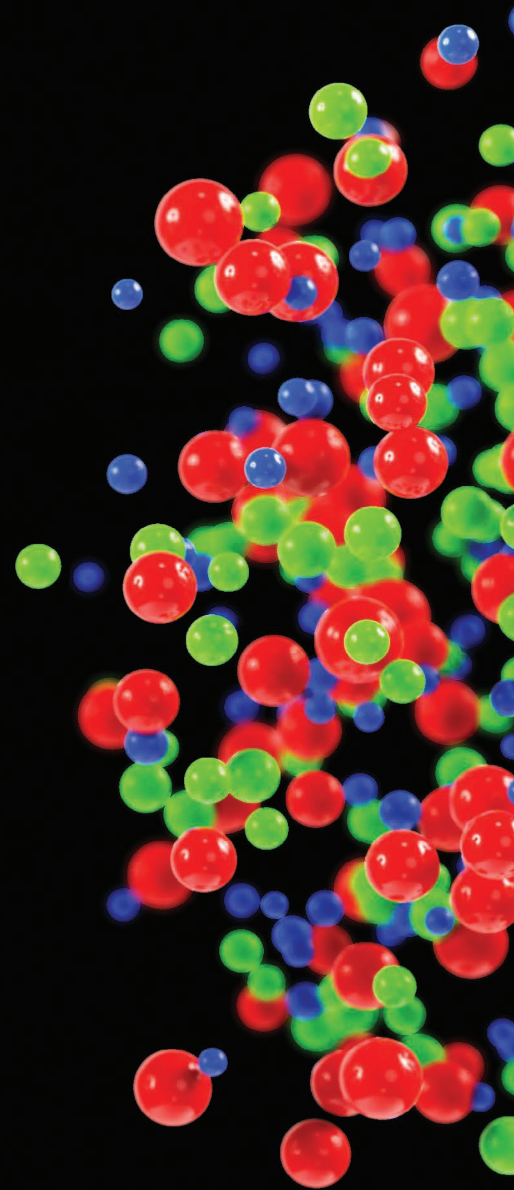


ILLUSTRATION BY Anatomy Blue

But do they? In bright sun, parts of an image are often washed out. In low light, images become grainy and unclear. Colors do not quite pop like those taken with a professional camera. And those are just the problems with cameras that record visible light. Although it would be great to have night vision in our cameras, infrared sensors cost a lot more for much poorer image quality than their visible-light brethren.

It's time for another revolution in imaging technology. This one will be brought to you by the quantum dot, a nanometer-size particle of semiconductor material, which acts much differently from its bulk counterpart.

**When a semiconductor** material absorbs light, it releases an electron from a chemical bond, and that electron is free to roam. The same process happens in a quantum dot (QD). But one thing is different: Although an electron is indeed released, it can't roam as easily; it gets squeezed by the edges of the particle, because the quantum dot is only a few nanometers in diameter. This squeeze is called quantum confinement, and it gives the particle some special properties.

The most useful property for imaging is that the light absorbed by the quantum dot is tunable—that is, the color can be continuously adjusted to almost any wavelength in the visible and infrared spectrum simply by choosing the right material and the right particle size. This tunability works in reverse as well—the color of the light emitted when the electron recombines can be selected precisely. It is this light-emission tunability that in recent years inspired the manufacturers of TVs and other kinds of displays to use quantum dots to improve color reproduction. (They've given the enhancement a number of names; the most common branding is "QLED.")

In addition to tunability, quantum dots have a few other nice features. Their small size allows these particles to be incorporated into printable inks, making quantum dots easy to slip into a manufacturing process. Quantum dots can absorb light more efficiently than silicon, which could allow camera makers to produce thinner image sensors. And QDs are sensitive across a broad dynamic range, from very low light to very high brightness.

**Before we tell you** how quantum-dot cameras will work—and when they will likely be commercially available—we should explain something about the CMOS sensor, today's state of the art for digital images. Clearly there has been considerable progress in the underlying technology in the past decade or two, particularly in making it smaller and cheaper. But the way in which it converts light into an image has largely remained unchanged.

In a typical camera, like the one in your phone, light passes through a series of lenses and a mosaic of red, green, and blue filters before being absorbed by one of the sensor pixels (sometimes called a photosite, to distinguish it from a pixel on an image) on the silicon CMOS chip. The filters determine which color each photosite will record.

When a photosite absorbs a photon, an electron is freed from a chemical bond and moves to an electrode at the edge of the pixel, where it is stored in a capacitor. A readout circuit converts the charge collected in each photosite over a set time to a voltage. The voltage determines the brightness for that pixel in the image.

A common manufacturing process creates both the silicon detectors and the readout circuits. This process involves a long but well-established series of steps of photolithography, etches, and growths. Such fabrication keeps costs low and is relatively simple. But it saddles silicon detectors with some disadvantages.

Typically, the readout electronics go on top of the detector, in what are called front-illuminated devices. Because of this placement, the metal contacts and traces reflect some of the incident light, decreasing efficiency. Back-illuminated devices avoid this reflection by having the readout electronics under the detector, but this placement increases fabrication cost and complexity. Only in the last decade has the cost of back-illuminated sensors dropped enough for them to be used in consumer devices, including phones and digital cameras.

Finally, silicon absorbs only wavelengths less than about 1 micrometer, so it won't work for imaging beyond the near-infrared range.

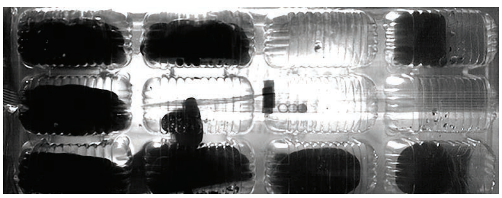
**Now let's look** at how quantum dots can change this equation.

As we mentioned before, by precisely tailoring the size of quantum dots, manufacturers of the materials can select exactly what wavelengths of light they absorb. The largest quantum dots in the visible spectrum, about 10 nanometers in diameter, absorb ultraviolet (UV), blue, and green light, and they emit red light, which is to say they're fluorescent. The smaller the QD, the more its absorption and emission shift toward blue in the color spectrum. For example, cadmium selenide QDs of about 3 nm absorb UV and blue light and emit green light.

Cameras with quantum-dot-based detectors operate basically the same way


SWIR VISION SYSTEMS

QUANTUM DOTS ON BOARD: The Acuros from SWIR Vision Systems [left] is the first commercially available infrared camera to use quantum-dot-based image sensors, giving it a cost advantage over traditional infrared cameras.



INFRARED VISION: Commercial applications of quantum-dot infrared cameras include quality control and food sorting. In these images, a standard CMOS visible-light sensor was used to record the full-color images; an Acuros infrared camera equipped with quantum-dot technology captured the monochrome images.

as their silicon CMOS counterparts. When a QD in a photosite absorbs a photon, an electron escapes its localized bond. The edge of the QD confines the electron's travels. However, if another QD is close enough, the free electron can "hop" over to it and, through sequential hops between QDs, reach the photosite's electrode where it can be counted by the pixel's readout circuit.

The readout circuits are manufactured in the same way as those built for silicon photodetectors—fabricated directly on a wafer. Adding the quantum dots to the wafer does add a processing step but an extremely simple one: They can be suspended in a solution as a sort of ink and printed or spin-coated over the circuitry.

Made in this way, quantum-dot photodetectors have the performance advantage of back-illuminated pixels, where nearly all the incident light reaches the detectors, without that technology's added cost and complexity.

And quantum dots have another advantage. Because they absorb light better than silicon, it takes only a thin layer atop the readout circuitry to gather almost all of the incoming photons, meaning the absorbing layer doesn't need to be nearly as thick as in standard CMOS image sensors. As a bonus, this thin, highly absorbing layer of QDs excels in both low light and high brightness, giving the sensor a better dynamic range.

**And, as Steve Jobs** used to say, "there's one more thing." Quantum-dot-based cameras have huge potential to bring infrared photography mainstream, because their tunability extends into infrared wavelengths.

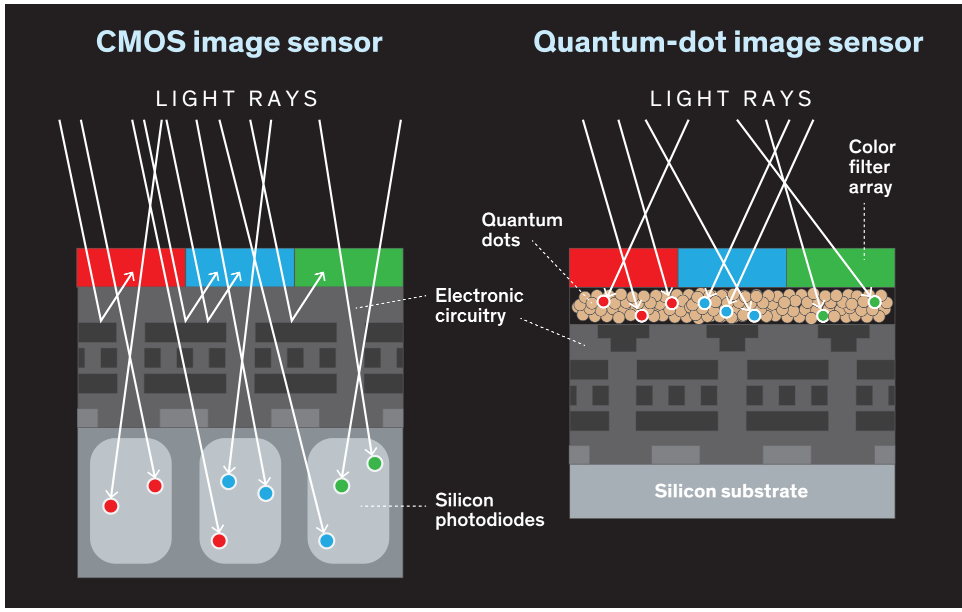
Today's infrared cameras function just like visible-light cameras, although the materials used for light absorption are quite different. Traditional infrared cameras use semiconductors with a small bandgap—such as lead selenide, indium antimonide, mercury cadmium telluride, or indium gallium arsenide—to absorb light that silicon does not. Pixel arrays made from these materials must be fabricated separately from the silicon CMOS circuits used to measure currents and generate an image. The detector array and circuit must then be connected at every pixel, typically by metal-to-metal bonding.

This time-consuming process, also known as hybridization, involves putting a small bump of low-melting-point indium on every pixel of both the detector array and the CMOS circuitry. The manufacturing machinery must then line the two up and press them together, then briefly melt the indium to create electrical connections. The complexity of this process limits the possible array sizes, pixel sizes, and sensor resolutions. Worse still, because it's done one camera chip at a time, hybridization is a low-throughput, costly process.

But quantum dots that are just as sensitive to infrared light as these traditional materials can be synthesized using inexpensive, large-scale chemical processing techniques. And, just as with

QD OR NOT QD: A quantum-dot image sensor for visible light [right], has several advantages over traditional CMOS technology [left], including its relative thinness, its elimination of reflections that prevent photons from being received, and the reduction of filtering errors caused by photons being received by the wrong photodiode.

their visible-light cousins, infrared-absorbing QDs can be painted onto chips after the silicon circuitry is complete, a quick and easy process needing no hybridization. Eliminating hybridization means that the resolution—the pixel size—can be less than the 15 µm or so needed to accommodate indium bumps, allowing for more pixels in a smaller area. A smaller sensor means smaller optics—and new shapes and sizes of infrared cameras at a far lower cost.




**All these factors** make quantum dots seem like a perfect imaging technology. But they aren't without challenges. Right now, the main obstacles to commercialization are stability, efficiency, and uniformity.

Manufacturers mainly solved these issues for the light-emitting quantum dots used in television displays by developing scalable chemical processes that enable the creation of high-efficiency dots in large quantities with very few defects. But quantum dots still oxidize in air, causing imperfections and changes to the sensor properties, including reduced sensitivity, increased noise, slower response time, and even shorting.

This stability problem didn't get in the way of commercialization of displays, however, because protecting the QDs used there from the atmosphere isn't terribly difficult. In the way that QDs are currently used in displays, the QD absorbs light from a blue LED and the photogenerated charge carriers stay within each individual quantum dot to recombine and fluoresce. So these QDs don't need to connect directly to circuitry, meaning that they can be protected by a surrounding polymer matrix with a barrier layer added on both sides of the polymer film, to prevent atmospheric exposure.

But for use in photodetection, sealing off individual QDs in a polymer won't work: The ejected electrons need to be free to migrate to the electrodes, where they can be counted.

One approach to allowing this migration while protecting the QDs from the ravages of the atmosphere would be to encapsulate the full layer of QDs or the entire device. That will likely be the initial solution. Alternatively, the QDs themselves could be specifically engineered to reduce the impact of oxidation without creating a barrier to charge transport, all while maintaining stability and processibility. Researchers are working toward that goal, but it's a tall order.

Another hurdle comes from the organic surfactants used today to maintain a stable solution of the quantum dots. These surfactants act as insulators, so they keep charge carriers from moving easily through the film of QDs to the electrode that collects the signal. Right now, manufacturers deal with this by depositing the QDs as a thin film and then replacing the long surfactant molecules with shorter ones that increase conductivity. But this adds a processing step and can make the QDs more susceptible to degrading over time, as the replacement process can damage the outer layer of QDs.

There is also a problem with the efficiency of photon detection. Due in part to their small size and large surface area, quantum dots can have many defects—imperfections in their crystal lattices that can cause photogenerated charges to recombine before the electron can reach an electrode. When this happens, the photon that initially hit the quantum dot is never detected by the circuitry, reducing the signal that ultimately reaches the camera's processor.

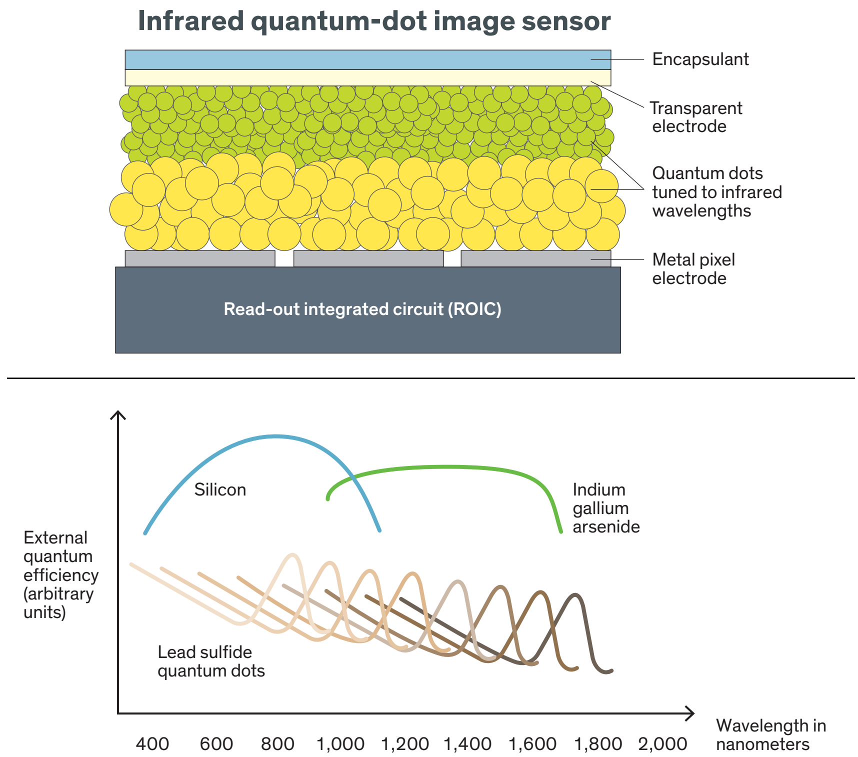
In traditional photodetectors—ones that contain single-crystal semiconductors—the defects are few and far between, resulting in efficiencies of greater than 50 percent. For QD-based photodetectors, this number is typically less than 20 percent. So in spite of the QDs themselves being better than silicon at absorb-

SOURCE: INVISAGE

ing light, the overall efficiency of QD-based photodetectors can't yet compete. But quantum-dot materials and device designs are improving steadily, with their efficiency continually getting better.

Because manufacturers use chemical processes to make quantum dots, there is some inherent variation in their size. And because the optical and electronic properties of a QD are driven by its size, any deviation from the desired diameter will cause a change in the color of light absorbed. With variations in the source chemicals, along with those in synthesis, purification, and storage, there can be significant size differences between one batch of QDs and another. The manufacturers must control their processes carefully to avoid this. Major companies with experience in this area have gotten quite good at maintaining uniformity, but smaller manufacturers often struggle to produce a consistent product.

WHY DOT? Manufacturability and tunability are among the advantages quantum-dot image sensors can claim over their CMOS counterparts. In the upper diagram, a cross section of SWIR Vision Systems' infrared image sensors shows a group of three pixels. The lower diagram shows the broad range of wavelengths of light that can be received by appropriately tuned quantum dots.




**In spite of** these challenges, companies have begun commercializing QD-based cameras, and these products are on the road to becoming mainstream.

A good early example is the Acuros camera, available from SWIR Vision Systems. That company is focused on manufacturing shortwave infrared quantum-dot cameras for use in applications where existing infrared cameras are too expensive. Its camera uses lead sulfide quantum dots, which absorb visible through shortwave infrared light. The detector in this camera currently has an average efficiency of 15 percent for infrared wavelengths, meaning that 15 percent of the photons that hit the detector end up as measurable signal. This is considerably lower than the efficiency of existing indium gallium arsenide technology, which can reach 80 percent. But with 15-µm pixels, the Acuros camera has a higher resolution than most infrared cameras. And it's sold at a price that, the company indicates, should be attractive to commercial users who cannot afford a traditional infrared camera—for applications like maritime imaging, produce inspection, and industrial-process monitoring.

As for the consumer camera market, in 2017 TechCrunch reported that Apple had acquired InVisage, a company dedicated to creating quantum-dot cameras for use in smartphones. Apple, as usual, has been quiet about its plans for this technology.

It may be that Apple is more interested in the infrared capabilities of QD-based cameras than their visible-light performance. Apple uses infrared light and sensors in its facial recognition technology, and cheaper chips with higher resolution for this purpose would clearly interest the company.

Other companies are also pushing hard to solve the stability and efficiency problems with quantum-dot photo sensors and to extend the boundaries of what is possible in terms of wavelength and sensitivity. BAE Systems, Brimrose, Episensors, and Voxtel are among those working to commercialize quantum-dot camera technology. Academic groups around the world are also deeply involved in QD-based sensor and camera research, including teams at MIT, University of Chicago, University of Toronto, ETH Zurich, Sorbonne University, and City University of Hong Kong.

Within five years, it's likely that we will have QD-based image sensors in our phones, enabling us to take better photos and videos in low light, improve facial recognition technology, and incorporate infrared photodetection into our daily lives in ways we can't yet predict. And they will do all of that with smaller sensors that cost less than anything available today. ■

**POST YOUR COMMENTS AT**
spectrum.ieee.org/qdcamera-mar2020

SOURCES: TOP: SWIR VISION SYSTEMS; BOTTOM: SENSORS, 2017

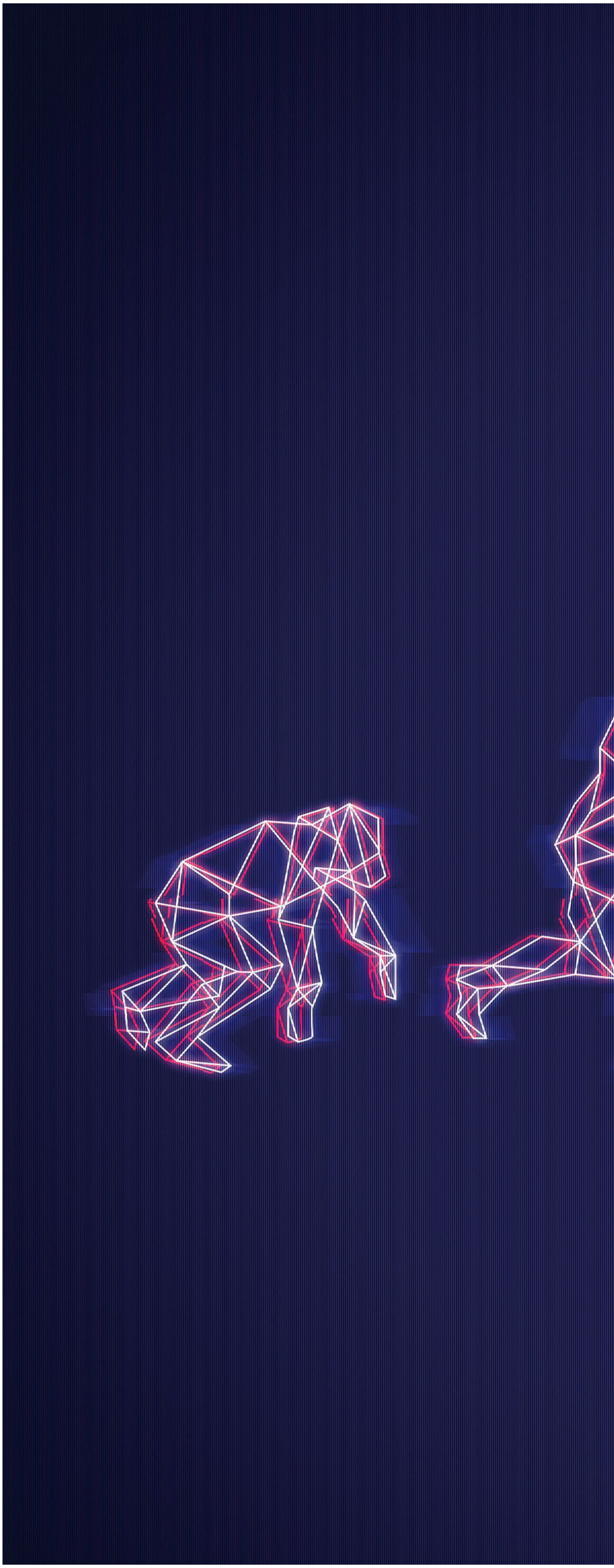
# AI

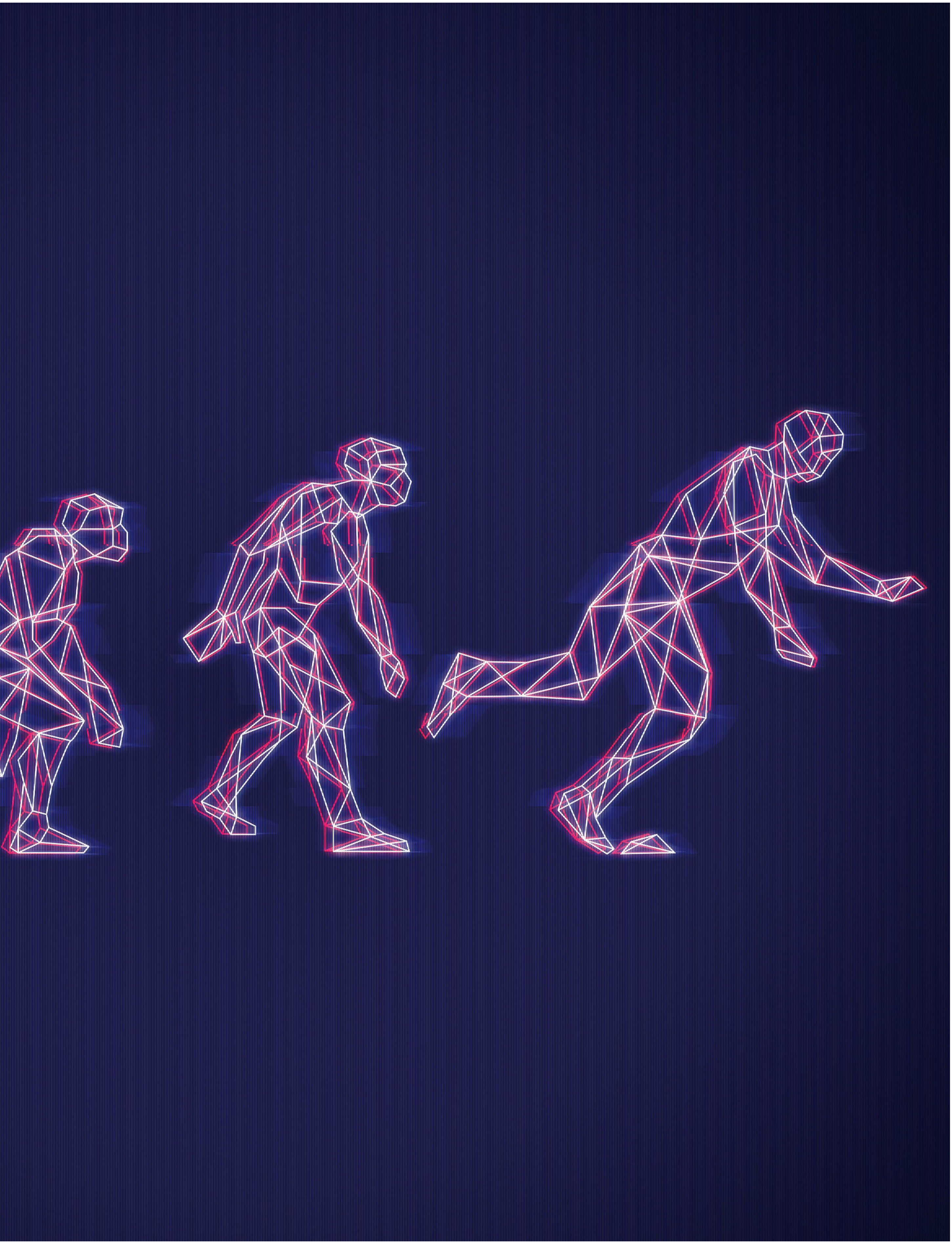
# Expect Evolution, Not Revolution

Despite the hype, artificial intelligence will take years to significantly boost economic productivity

By Jeffrey Funk

Illustrations by Edmon de Haro





**In 2016, London-based DeepMind Technologies, a subsidiary of Alphabet** (which is also the parent company of Google), startled industry watchers when it reported that the application of artificial intelligence had reduced the cooling bill at a Google data center by a whopping 40 percent. What's more, we learned that year, DeepMind was starting to work with National Grid in the United Kingdom to save energy throughout the country using deep learning to optimize the flow of electricity. ● Could AI really slash energy usage so profoundly? In the years that have since passed, I've searched for articles on the application of AI to other data centers but find no evidence of important gains. What's more, DeepMind's talks with National Grid about energy have broken down. And the financial results for DeepMind certainly don't suggest that customers are lining up for its services: For 2018, the company reported losses of US $571 million on revenues of $125 million, up from losses of $366 million in 2017. Last April, *The Economist* characterized DeepMind's 2016 announcement as a publicity stunt, quoting one inside source as saying, "[DeepMind just wants] to have some PR so they can claim some value added within Alphabet."

This episode encouraged me to look more deeply into the economic promise of AI and the rosy projections made by champions of this technology within the financial sector. This investigation was just the latest twist on a long-standing interest of mine. In the early 1980s, I wrote a doctoral dissertation on the economics of robotics and AI, and throughout my career as a professor and technology consultant I have followed the economic projections for AI, including detailed assessments by consulting organizations such as Accenture, PricewaterhouseCoopers International (PwC), and McKinsey & Co.

These analysts have lately been asserting that AI-enabled technologies will dramatically increase economic output. Accenture claims that by 2035 AI will double growth rates for 12 developed countries and increase labor productivity by as much as a third. PwC claims that AI will add $15.7 trillion to the global economy by 2030, while McKinsey projects a $13 trillion boost by that time.

Other forecasts have focused on specific sectors such as retail, energy, education, and manufacturing. In particular, the McKinsey Global Institute assessed the impact of AI on these four sectors in a 2017 report titled *Artificial Intelligence: The New Digital Frontier?* and did so for a much longer list of sectors in a 2018 report. In the latter, the institute concluded that AI techniques "have the potential to create between $3.5 trillion and $5.8 trillion in value annually across nine business functions in 19 industries. This constitutes about 40 percent of the overall $9.5 trillion to $15.4 trillion annual impact that could potentially be enabled by all analytical techniques."

Wow. These are big numbers. If true, they create a powerful incentive for companies to pursue AI—with or without help from McKinsey consultants. But are these predictions really valid?

Many of McKinsey's estimates were made by extrapolating from claims made by various startups. For instance, its prediction of a 10 percent improvement in energy efficiency in the U.K. and elsewhere was based on the purported success of DeepMind and also of Nest Labs, which became part of Google's hardware division in 2018. In 2017, Nest, which makes a smart thermostat and other intelligent products for the home, lost $621 million on revenues of $726 million. That fact doesn't mesh with the notion that Nest and similar companies are contributing, or are poised to contribute, hugely to the world economy.

So I decided to investigate more systematically how well such AI startups were doing. I found that many were proving not nearly as valuable to society as all the hype would suggest. This assertion will certainly rub a lot of people the wrong way, the analysts at McKinsey among them. So I'd like to describe here how I reached my much more pessimistic conclusions.

**My investigation of Nest Labs** expanded into a search for evidence that smart meters in general are leading to large gains in energy efficiency. In 2016, the British government began a coordinated campaign to install smart meters throughout the country by 2020. And since 2010, the U.S. Department of Energy has invested some $4.5 billion installing more than 15 million smart meters throughout the United States. Curiously enough, all that effort has had little observed impact on energy usage. The U.K. government recently revised downward the amount it figures a smart meter will save each household annually, from £26 to just £11. And the cost of smart meters and their installation has risen, warns the U.K.'s National Audit Office. All of this is not good news for startups banking on the notion that smart thermostats, smart home appliances, and smart meters will lead to great energy savings.

## AI's TOP 40

The author analyzed 40 U.S.-based AI startups with valuations greater than US $1 billion or with more than $70 million in equity funding. They are shown here grouped according to the areas of technology they address.

### BASIC COMPUTER HARDWARE OR SOFTWARE

Tanium Cybereason Sentient Technologies
CloudMinds Uptake Technologies
Cylance Wave Computing Dataminr
DataRobot OpenAI Petuum H2O
CrowdStrike Shape Security Ayasdi
Endgame Trifacta

### HEALTH CARE

Flatiron Health
Tempus Labs
Freenome

### FINANCE

Avant
ZestFinance
Upstart

### AUTOMATION TOOLS

Automation Anywhere
UiPath Xant Brain Corp.
ZipRecruiter WorkFusion
Conversica Algolia

### TRANSPORTATION

Zoox Nauto Nuro

### BIO/AGRO

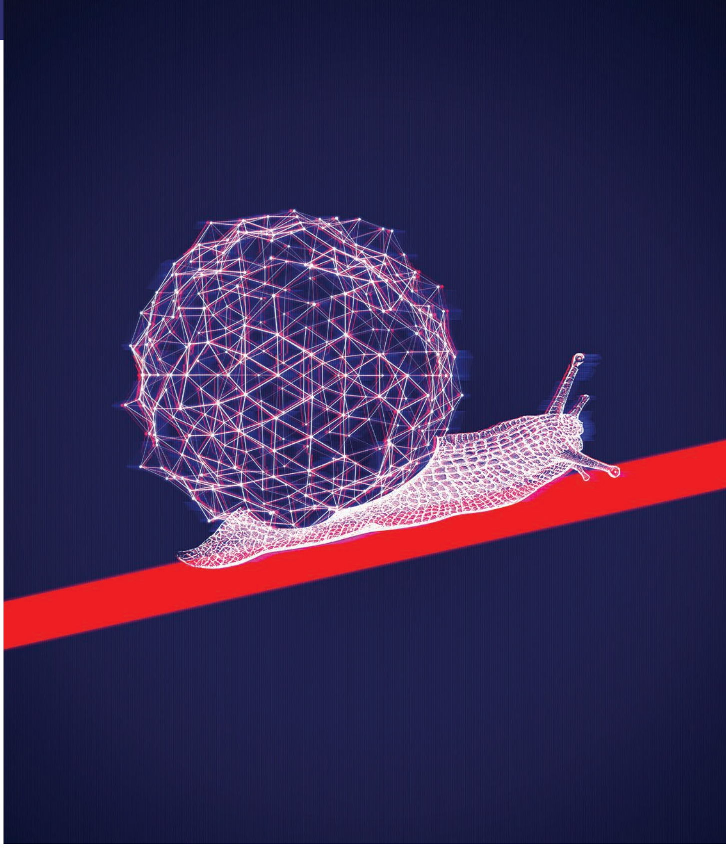
Indigo Agriculture
Zymergen

### OTHER

Afiniti Vicarious
SoundHound Orbital Insight

**KEY** (valuation or funding)

Under US $100 million | $100 million to $500 million | $500 million to $1 billion | Over $1 billion



Are other kinds of AI startups having a greater positive effect on the economy? Tech sector analyst CB Insights reports that overall venture capital funding in the United States was $115 billion in 2018, of which $9.3 billion went to AI startups. While that's just 8 percent of the total, it's still a lot of money, indicating that there are many U.S. startups working on AI (although some overstate the role of AI in their business plans to acquire funding).

To probe further, I gathered data on the U.S. AI startups that have received the most funding and looked at which industries they were hoping to disrupt. The reason for focusing on the United States is that it has the longest history of startup success, so it seems likely that its AI startups are more apt to flourish than those in other countries. My intention was to evaluate whether these U.S. startups had succeeded in shaking up various industries and boosting productivity, or whether they promise to do so shortly.

In all, I examined 40 U.S. startups working on AI. These either had valuations greater than $1 billion or had more than $70 million in equity funding. Other than two that had been acquired by public companies, the startups I looked at are all private firms. I found their names and product offerings in lists of leading startups that Crunchbase, *Fortune*, and Datamation had compiled and published. I then updated my data set with more recent news about these companies (including reports of some shutdowns).

I categorized these 40 startups by the type of product or service they offered. Seventeen are working on what I would call basic computer hardware and software (Wave Computing and OpenAI, respectively, are examples), including cybersecurity (CrowdStrike, for instance). That is, I included in this category companies building tools that are intended to support the computing environment itself.

Making up another large fraction—8 of the 40—are companies that develop software that automates various tasks. The robotic process-automation software being developed by Automation Anywhere, UiPath, and WorkFusion, for example, enables higher productivity among white-collar workers. Software from Brain Corp. converts manual equipment into intelligent robots. Algolia, Conversica, and Xant offer software to improve sales and marketing. ZipRecruiter targets human resources.

The remaining startups on my list are spread among various industries. Three (Flatiron Health, Freenome, Tempus Labs) work in health care; three more (Avant, Upstart, ZestFinance) are focused on financial technology; two (Indigo Agriculture, Zymergen) target agriculture or synthetic biology; and three others (Nauto, Nuro, Zoox) involve transportation. There is just one startup each for geospatial analytics (Orbital Insight), patterns of human interaction (Afiniti), photo/video recognition (Vicarious), and music recognition (SoundHound).

Are there indications that these startups will bring large productivity improvements in the near future? In my view, software that automates tasks normally carried out by white-collar workers is probably the most promising of the products and services that AI is being applied to. Similar to past improvements in tools for white-collar professionals, including Excel for accountants and computer-aided design for engineers and architects, these types of AI-based tools have the greatest potential impact on productivity. For instance, there are high hopes for generative design, in which teams of people input constraints and the system proposes specific designs.

But looking at the eight startups on my list that are working on automation tools for white-collar workers, I realized that they are not targeting things that would lead to much higher productivity. Three of them are focused on sales and marketing, which is often a zero-sum game: The company with the best software takes customers from competitors, with only small increases in productivity under certain conditions. Another one of these eight companies is working on human-resource software, whose productivity benefits may be larger than those for sales and marketing but probably not as large as you'd get from improved robotic process automation.

This leaves four startups that do offer such software, which may lead to higher productivity and lower costs. But even among these startups, none currently offers software that helps engineers and architects become more productive through, for example, generative design. Software of this kind isn't coming from the largest startups, perhaps because there is a strong incumbent, Autodesk, or because the relevant AI is still not developed enough to provide truly useful tools in this area.

The relatively large number of startups I classified as working on basic hardware and software for computing (17) also suggests that productivity improvements are still many years away. Although basic hardware and software are a necessary part of developing higher-level AI-based tools, particularly ones utilizing machine learning, it will take time for the former to enable the latter. I suppose this situation simply reflects that AI is still in its infancy. You certainly get that impression from companies like OpenAI: Although it has received $1 billion in funding (and a great deal of attention), the vagueness of its mission—"Benefiting all of humanity"—suggests that it will take many years yet for specific useful products and services to evolve from this company's research.

The large number of these startups that are focused on cybersecurity (seven) highlights the increasing threat of security problems, which raise the cost of doing business over the Internet. AI's ability to address cybersecurity issues will likely make the Internet more safe, secure, and useful. But in the end, this thrust reflects yet higher costs in the future for Internet businesses and will not, to my mind, lead to large productivity improvements within the economy as a whole.

If not from the better software tools it brings, where will AI bring substantial economic gains? Health care, you would think, might benefit greatly from AI. Yet the number of startups on my list that are applying AI to health care (three) seems oddly small if that were really the case. Perhaps this has something to do with IBM's experience with its Watson AI, which proved a disappointment when it was applied to medicine.

Still, many people remain hopeful that AI-fueled health care startups will fill the gap left by Watson's failures. Arguing against this is Robert Wachter, who points out that it's much more difficult to apply computers to health care than to other sectors. His 2015 book, *The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age* (McGraw-Hill Education), details the many reasons that health care lags other industries in the application of computers and software. It's not clear that adding AI to the mix of digital technologies available will do anything to change the situation.

There are also some big applications missing from the list of well-funded AI startups. Housing represents the largest category of consumer expenditures in the United States, but none of these startups are addressing this sector of the economy at all. Transportation is the second largest expenditure, and it is the focus of just three of these startups. One is working on a product that identifies distracted drivers. Another intends to provide automated local deliveries. Only one startup on the list is developing driverless passenger vehicles. That there is only one working on self-driving cars is consistent with the pessimism recently expressed by executives of Ford, General Motors, and Mercedes-Benz about the prospects for driverless vehicles taking to the streets in large numbers anytime soon, even though $35 billion has already been spent on R&D for them.

**Admittedly, my assessment of** what these 40 companies are doing and whether their offerings will shake up the world over the next decade is subjective. Perhaps it makes better sense to consider a more objective measure of whether these companies are providing value to the world economy: their profitability.

Alas, good financial data is not available on privately held startups, and only two of the companies on my list are now part of public companies. Also, startups often take years to turn a profit (Amazon took seven years). So there isn't a lot to go on here. Still, there are some broad trends in the tech sector that are quite telling.

The fraction of tech companies that are profitable by the time they go public dropped from 76 percent in 1980 to just 17 percent in 2018, even though the average time to IPO has been rising—it went from 2.8 years in 1998 to 7.7 years in 2016, for example. Also, the losses of some well-known startups that took a long time to go public are huge. For instance, none of the big ride-sharing companies are making a profit, including those in the United States (Uber and Lyft), China, India, and Singapore, with total losses of about $5 billion in 2018. Most bicycle and scooter sharing, office sharing, food delivery, P2P (peer-to peer) lending, health care insurance and analysis, and other consumer service startups are also losing vast amounts of money, not only in the United States but also in China and India.

Most of the 40 AI startups I examined will probably stay private, at least in the near term. But even if some do go public several years down the road, it's unlikely they'll be profitable at that point, if the experience of many other tech companies is any guide. It may take these companies years more to achieve the distinction of making more money than they are spending.

**For the reasons I've given,** it's very hard for me to feel confident that any of the AI startups I examined will provide the U.S. economy with a big boost over the next decade. Similar pessimism is also starting to emerge from such normally cheery publications as *Technology Review* and *Scientific American*. Even the AI community is beginning to express concerns in books such as *The AI Delusion* and *Rebooting AI: Building Artificial Intelligence We Can Trust*, concerns that are growing amid the rising hype about many new technologies.

The most promising areas for rapid gains in productivity are likely to be found in robotic process automation for white-collar workers, continuing a trend that has existed for decades. But these improvements will be gradual, just as those for computer-aided design and computer-aided engineering software, spreadsheets, and word processing have been.

Viewed over the span of decades, the value of such software is impressive, bringing huge gains in productivity for engineers, accountants, lawyers, architects, journalists, and others—gains that enabled some of these professionals (particularly engineers) to enrich the global economy in countless ways.

Such advances will no doubt continue with the aid of machine learning and other forms of AI. But they are unlikely to be nearly as disruptive—for companies, for workers, or for the economy as a whole—as many observers have been arguing. ■

**POST YOUR COMMENTS AT** spectrum.ieee.org/aistartups-mar2020

By Shaoshan Liu & Jean-Luc Gaudiot

# Autonomous Vehicles Lite

## Self-driving technologies should start small, go slow

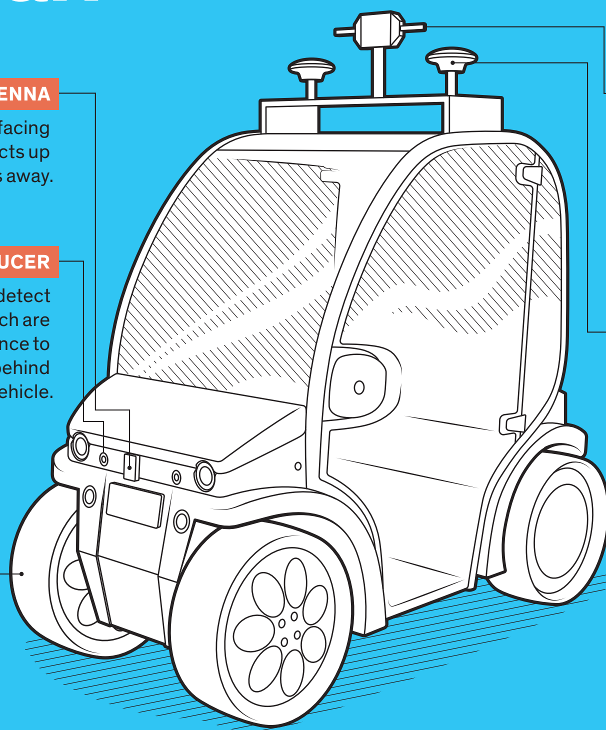
**Many young urbanites** don't want to own a car, and unlike earlier generations, they don't have to rely on mass transit. Instead they treat mobility as a service: When they need to travel significant distances, say, more than 5 miles (8 kilometers), they use their phones to summon an Uber (or a car from a similar ride-sharing company). If they have less than a mile or so to go, they either walk or use various "micromobility" services, such as the increasingly ubiquitous Lime and Bird scooters or, in some cities, bike sharing. • The problem is that today's mobility-as-a-service ecosystem often doesn't do a good job covering intermediate distances, say a few miles. Hiring an Uber or Lyft for such short trips proves frustratingly expensive, and riding a scooter or bike more than a mile or so can be taxing to many people. So getting yourself to a destination that is from 1 to 5 miles away can be a challenge. Yet such trips account for about half of the total passenger miles traveled. • Many of these intermediate-distance trips take place in environments with limited traffic, such as university campuses and industrial parks, where it is now both economically reasonable and technologically possible to deploy small, low-speed autonomous vehicles powered by electricity. We've been involved with a startup that intends to make this form of transportation popular. The company, PerceptIn, has



ILLUSTRATION BY L-Dopa

# Look Ma, No Lidar!

Although lidar is de rigueur for highway-capable autonomous cars, low-speed vehicles can drive themselves using a suite of less sophisticated (and less costly) sensors.




autonomous vehicles operating at tourist sites in Nara and Fukuoka, Japan; at an industrial park in Shenzhen, China; and is just now arranging for its vehicles to shuttle people around Fishers, Ind., the location of the company's headquarters.

Because these diminutive autonomous vehicles never exceed 20 miles (32 kilometers) per hour and don't mix with high-speed traffic, they don't engender the same kind of safety concerns that arise with autonomous cars that travel on regular roads and highways. While autonomous driving is a complicated endeavor, the real challenge for PerceptIn was not about making a vehicle that can drive itself in such environments—the technology to do that is now well established—but rather about keeping costs down.

Given how expensive autonomous cars still are in the quantities that they are currently being produced—an experimental model can cost you in the neighborhood of US $300,000—you might not think it possible to sell a self-driving vehicle of any kind for much less. Our experience over the past few years shows that, in fact, it is possible today to produce a self-driving passenger vehicle much more economically: PerceptIn's vehicles currently sell for about $70,000, and the price will surely drop in the future. Here's how we and our colleagues at PerceptIn brought the cost of autonomous driving down to earth.

**Let's start by explaining** why autonomous cars are normally so expensive. In a nutshell, it's because the sensors and computers they carry are very pricey.

The suite of sensors required for autonomous driving normally includes a high-end satellite-navigation receiver, lidar (light detection and ranging), one or more video cameras, radar, and sonar. The vehicle also requires at least one very powerful computer.

The satellite-navigation receivers used in this context aren't the same as the one found in your phone. The kind built into autonomous vehicles have what is called real-time kinematic capabilities for high-precision position fixes—down to 10 centimeters. These devices typically cost about $4,000. Even so, such satellite-navigation receivers can't be entirely relied on to tell the vehicle where it is. The fixes it gets could be off in situations where the satellite signals bounce off of nearby buildings, introducing noise and delays. In any case, satellite navigation requires an unobstructed view of the sky. In closed environments, such as tunnels, that just doesn't work.

Fortunately, autonomous vehicles have other ways to figure out where they are. In particular they can use lidar, which determines distances to things by bouncing a laser beam off them and measuring how long it takes for the light to reflect back. A typical lidar unit for autonomous vehicles covers a range of 150 meters and samples more than 1 million spatial points per second.

Such lidar scans can be used to identify different shapes in the local environment.



ILLUSTRATION BY Chris Philpot

The vehicle's computer then compares the observed shapes with the shapes recorded in a high-definition digital map of the area, allowing it to track the exact position of the vehicle at all times. Lidar can also be used to identify and avoid transient obstacles, such as pedestrians and other cars.

Lidar is a wonderful technology, but it suffers from two problems. First, these units are extremely expensive: A high-end lidar for autonomous driving can easily cost more than $80,000, although costs are dropping, and for low-speed applications a suitable unit can be purchased for about $4,000. Also, lidar, being an optical device, can fail to provide reasonable measurements in bad weather, such as heavy rain or fog.

The same is true for the cameras found on these vehicles, which are mostly used to recognize and track different objects, such as the boundaries of driving lanes, traffic lights, and pedestrians. Usually, multiple cameras are mounted around the vehicle. These cameras typically run at 60 frames per second, and the multiple cameras used can generate more than 1 gigabyte of raw data each second. Processing this vast amount of information, of course, places very large computational demands on the vehicle's computer. On the plus side, cameras aren't very expensive.

The radar and sonar systems found in autonomous vehicles are used for obstacle avoidance. The data sets they generate show the distance from the nearest object in the vehicle's path. The major advantage of these systems is that they work in all weather conditions. Sonar usually covers a range of up to 10 meters, whereas radar typically has a range of up to 200 meters. Like cameras, these sensors are relatively inexpensive, often costing less than $1,000 each.

The many measurements such sensors supply are fed into the vehicle's computers, which have to integrate all this information to produce an understanding of the environment. Artificial neural networks and deep learning, an approach that has grown rapidly in recent years, play a large role here. With these techniques, the computer can keep track of other vehicles moving nearby, as well as of pedestrians crossing the road, ensuring the autonomous vehicle doesn't collide with anything or anyone.

Of course, the computers that direct autonomous vehicles have to do a lot more than just avoid hitting something. They have to make a vast number of decisions about where to steer and how fast to go. For that, the vehicle's computers generate predictions about the upcoming movement of nearby vehicles before deciding on an action plan based on those predictions and on where the occupant needs to go.

Lastly, an autonomous vehicle needs a good map. Traditional digital maps are usually generated from satellite imagery and have meter-level accuracy. Although that's more than sufficient for human drivers, autonomous vehicles demand higher accuracy for lane-level information. Therefore, special high-definition maps are needed.

**SLOWLY BUT SURELY:** The authors' approach to autonomy has been applied to two different types of low-speed electric vehicles. One is a two-seat "pod," shown here being demonstrated at Purdue University, where it was used to transport students from parking lots to the center of campus [top]. The other is a multipassenger bus, which is being used now at various sites around the world, including the Nara Palace historical park in Japan [bottom].

PERCEPTIN (2)

Just like traditional digital maps, these HD maps contain many layers of information. The bottom layer is a map with grid cells that are about 5 by 5 cm; it's generated from raw lidar data collected using special cars. This grid records elevation and reflection information about the objects in the environment.

On top of that base grid, there are several layers of additional information. For instance, lane information is added to the grid map to allow autonomous vehicles to determine whether they are in the correct lane. On top of the lane information, traffic-sign labels are added to notify the autonomous vehicles of the local speed limit, whether they are approaching traffic lights, and so forth. This helps in cases where cameras on the vehicle are unable to read the signs.

Traditional digital maps are updated every 6 to 12 months. To make sure the maps that autonomous vehicles use contain up-to-date information, HD maps should be refreshed weekly. As a result, generating and maintaining HD maps can cost millions of dollars per year for a midsize city.

All that data on those HD maps has to be stored on board the vehicle in solid-state memory for ready access, adding to the cost of the computing hardware, which needs to be quite powerful. To give you a sense, an early computing system that Baidu employed for autonomous driving used an Intel Xeon E5 processor and four to eight Nvidia K80 GPU accelerators. The system was capable of delivering 64.5 trillion floating-point operations per second, but it consumed around 3,000 watts and generated an enormous amount of heat. And it cost about $30,000.

**Given that the sensors** and computers alone can easily cost more than $100,000, it's not hard to understand why autonomous vehicles are so expensive, at least today. Sure, the price will come down as the total number manufactured increases. But it's still unclear how the costs of creating and maintaining HD maps will be passed along. In any case, it will take time for better technology to address all the obvious safety concerns that come with | CONTINUED ON PAGE 48

# BIG IN JAPAN

HOW ART, TECH, AND PEPSICO COLLABORATED, THEN CLASHED AT THE 1970 WORLD'S FAIR **BY W. PATRICK McCRAY**

**ON 18 MARCH 1970,** a former Japanese princess stood at the center of a cavernous domed structure on the outskirts of Osaka. With a small crowd of dignitaries, artists, engineers, and business executives looking on, she gracefully cut a ribbon that tethered a large red balloon to a ceremonial Shinto altar. Rumbles of thunder rolled out from speakers hidden in the ceiling. As the balloon slowly floated upward, it appeared to meet itself in midair, reflecting off the massive spherical mirror that covered the walls and ceiling.

With that, one of the world's most extravagant and expensive multimedia installations officially opened, and the attendees turned to congratulate one another on this collaborative melding of art, science, and technology. Underwritten by PepsiCo, the installation was the beverage company's signal contribution to Expo '70, the first international exposition to be held in an Asian country.

A year and a half in the making, the Pepsi Pavilion drew eager crowds and elicited effusive reviews. And no wonder: The pavilion was the creation of Experiments in Art and Technology–E.A.T.–an influential collective of artists, engineers, technicians, and scientists based in New York City. Led by Johan Wilhelm "Billy" Klüver, an electrical engineer at Bell Telephone Laboratories, E.A.T. at its peak had more than a thousand members and enjoyed generous support from corporate donors and philanthropic foundations. Starting in the mid-1960s and continuing into the '70s, the group mounted performances and installations that blended electronics, lasers, telecommunications, and computers with artistic interpretations of current events, the natural world, and the human condition.

E.A.T. members saw their activities transcending the making of art. Artist-engineer collaborations were understood as creative experiments that would benefit not just the art world but also industry and academia. For engineers, subject to vociferous attacks about their complicity in the arms race, the Vietnam War, environmental destruction, and other global ills, the art-and-technology movement presented an opportunity to humanize their work.

Accordingly, Klüver and the scores of E.A.T. members in the United States and Japan who designed and built the pavilion considered it an "experiment in the scientific sense," as the 1972 book *Pavilion: Experiments in Art and Technology* stated. Klüver pitched the installation as a "piece of hardware" that engineers and artists would program with "software" (that is, live performances) to create an immersive visual, audio, and tactile experience. As with other E.A.T. projects, the goal was not about the *product* but the *process*.

Pepsi executives, unsurprisingly, viewed their pavilion on somewhat different terms. These were the years of the Pepsi Generation, the company's mildly countercultural branding. For them, the pavilion would be at once an advertisement, a striking visual statement, and a chance to burnish the company's global reputation. To that end, Pepsi directed close to US $2 million (over $13 million today) to E.A.T. to create the biggest, most elaborate, and most expensive art project of its time.

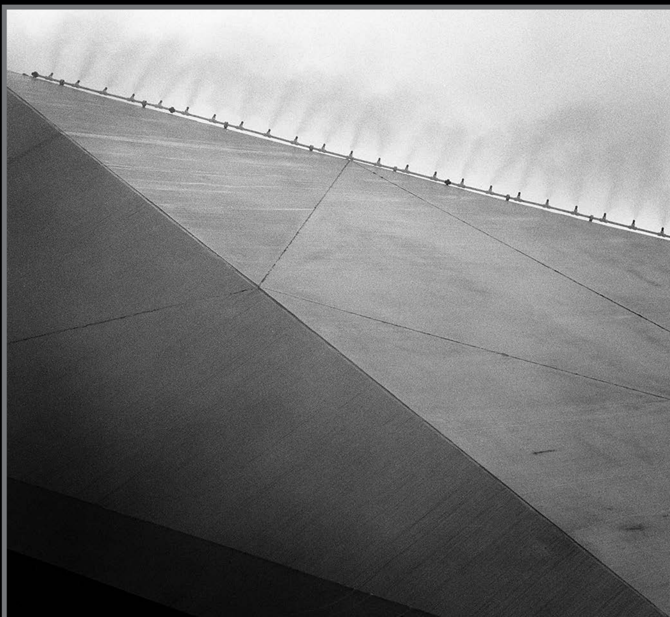
Perhaps it was inevitable, but over the 18 months it took E.A.T. to execute the project, Pepsi executives grew increasingly concerned about the group's vision. Just a month after the opening, the partnership collapsed amidst a flurry of recriminating letters and legal threats. And yet, despite this inglorious end, the participants considered the pavilion a triumph.

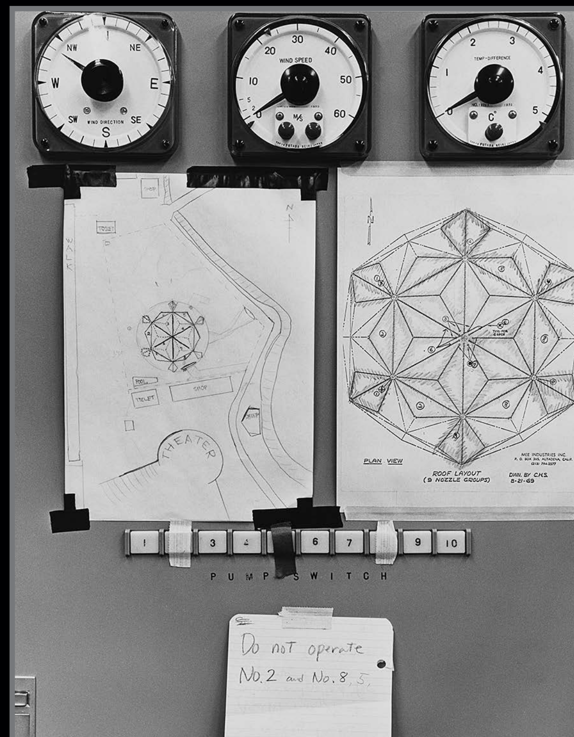
**The pavilion was born** during a backyard conversation in the fall of 1968 between David Thomas, vice president in charge of Pepsi's marketing, and his neighbor, Robert Breer, a sculptor and filmmaker who belonged to the E.A.T. collective. Pepsi had planned to contract with Disney to build its Expo '70 exhibition, as it had done for the 1964 World's Fair in New York City. Some Pepsi executives were, however, concerned that the conservative entertainment company wouldn't produce something hip enough for the burgeoning youth market, and they had memories of the 1964 project, when Disney ran well over its already considerable budget. Breer put Thomas in touch with Klüver, productive dialogue ensued, and the company hired E.A.T. in December 1968.

Klüver was a master at straddling the two worlds of art and science. Born in Monaco in 1927 and raised in Stockholm, he developed a deep appreciation for cinema as a teen, an interest he maintained while studying with future Nobel physicist Hannes Alfvén. After earning a Ph.D. in electrical engineering at the University of California, Berkeley, in 1957, he accepted a coveted research position at Bell Labs in Murray Hill, N.J.

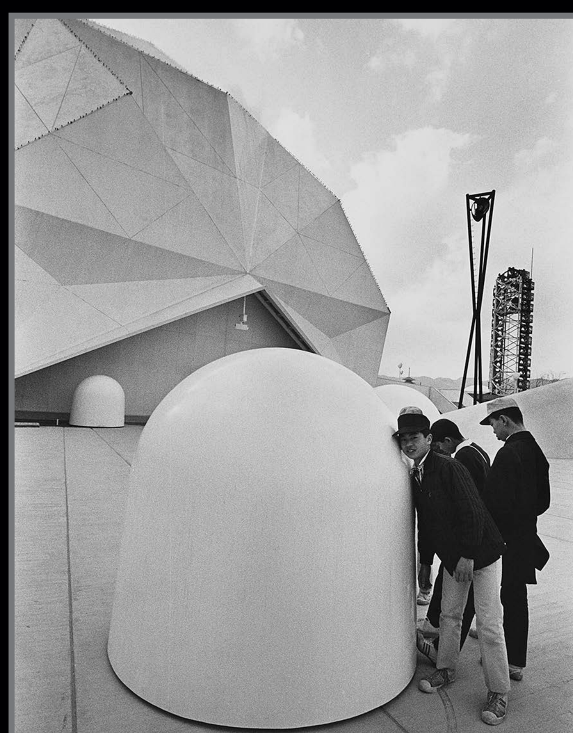
While keeping up a busy research program, Klüver made time to explore performances and gallery openings in downtown Manhattan and to seek out artists. He soon began collaborating with artists such as Yvonne Rainer, Andy Warhol,

# I. The Fog and The Floats



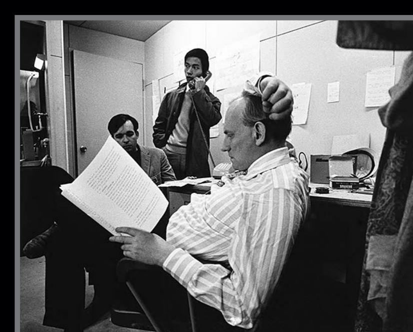



**ART MEETS TECH:** Artist Fujiko Nakaya and physicist Thomas R. Mee created artificial fog by spraying pure water through narrow nozzles installed on the Pepsi Pavilion's roof [above left]. The system tracked wind speed and direction [above right] to ensure the fog was distributed over the building's surface. On the pavilion's terrace, autonomous white "floats" built by sculptor Robert Breer [kneeling, below left] roamed about, emitting soft sounds [below right].



# II. The Experimenters

**WORK IN PROGRESS: The Pepsi Pavilion [below left] was overseen by Bell Labs engineer Billy Klüver [below right], who called it "an experiment in the scientific sense." Dozens of artists and engineers in the United States and Japan worked on the project, including [at right] Elsa Garmire [hoop earrings], Thomas R. Mee [moustache], and Fujiko Nakaya [white turtleneck]. The pavilion's elaborate audio system was designed by David Tudor [below center].**



Jasper Johns, and Robert Rauschenberg, contributing his technical expertise and helping to organize exhibitions and shows. His collaboration with Jean Tinguely on a self-destructing sculpture, called *Homage to New York*, appeared on the April 1969 cover of *IEEE Spectrum*. Klüver emerged as the era's most visible and vocal spokesperson for the merger of art and technology in the United States. *Life* magazine called him the "Edison-Tesla-Steinmetz-Marconi-Leonardo da Vinci of the American avant-garde."

Klüver's supervisor, John R. Pierce, was tolerant and even encouraging of his activities. Pierce had his own creative bent, writing science fiction in his spare time and collaborating with fellow Bell engineer Max Mathews to create computer-generated music. Meanwhile, Bell Labs, buoyed by the economic prosperity of the 1960s, supported a small coterie of artists-in-residence, including Nam June Paik, Lillian Schwartz, and Stan VanDerBeek.

In time, Klüver devised more ambitious projects. For his 1966 orchestration of *9 Evenings: Theatre and Engineering*, nearly three dozen engineering colleagues worked with artists to build wireless radio transmitters, carts that floated on cushions of air, an infrared television system, and other electronics. Held at New York City's 69th Regiment Armory—which in 1913 had hosted a pathbreaking exhibition of modern art—*9 Evenings* expressed a new creative culture in which artists and engineers collaborated.

In the midst of organizing *9 Evenings*, Klüver, along with artists Rauschenberg

SHUNK-KENDER/J. PAUL GETTY TRUST/GETTY RESEARCH INSTITUTE, LOS ANGELES, 2014.R.20 (4) and Robert Whitman and Bell Labs engineer Fred Waldhauer, founded Experiments in Art and Technology. By the end of 1967, more than a thousand artists and technical experts had joined. And a year later, E.A.T. had scored the commission to create the Pepsi Pavilion.

**From the start,** E.A.T. envisioned the pavilion as a multimedia environment that would offer a flexible, personalized experience for each visitor and that would express irreverent, uncommercial, and antiauthoritarian values.

But reaching consensus on how to realize that vision took months of debate and argument. Breer wanted to include his slow-moving cybernetic “floats”–large, rounded, self-driving sculptures powered by car batteries. Whitman was becoming intrigued with lasers and visual perception, and felt there should be a place for that. Forrest “Frosty” Myers argued for an outdoor light installation using searchlights, his focus at the time. Experimental composer David Tudor imagined a sophisticated sound system that would transform the Pepsi Pavilion into both recording studio and instrument.

“We’re all painters,” Klüver recalled Rauschenberg saying, “so let’s do something nonpainterly.” Rauschenberg’s attempt to break the stalemate prompted a further flood of suggestions. How about creating areas where the temperature changed? Or pods that functioned as anechoic chambers–small spaces of total silence? Maybe the floor could have rear-screen projections that gave visitors the impression of walking over flames, clouds, or swimming fish. Perhaps wind tunnels and waterfalls could surround the entrances.

Eventually, Klüver herded his fellow E.A.T. members into agreeing to an eclectic set of tech-driven pieces. The pavilion building itself was a white, elongated geodesic dome, which E.A.T. detested and did its best to obscure. And so a visitor approaching the finished pavilion encountered not the building but a veil of artificial fog that completely enshrouded the structure. At night, the fog was dramatically lit and framed by high-intensity xenon lights designed by Myers.

On the outdoor terrace, Breer’s white floats rolled about autonomously like large bubbles, emitting soft sounds–speech, music, the sound of sawing wood–and gently reversing themselves when they bumped into something. Steps led downward into a darkened tunnel, where visitors were greeted by a Japanese hostess wearing a futuristic red dress and bell-shaped hat and handed a clear plastic wireless handset. Stepping farther into the tunnel, they would be showered with red, green, yellow, and blue light patterns from a krypton laser system, courtesy of Whitman.

Ascending into the main pavilion, the visitors’ attention would be drawn immediately upward, where their reflections off the huge spherical mirror made it appear that they were floating in space. The dome also created auditory illusions, as echoes and reverberations toyed with people’s sense of acoustic reality. The floors of the circular room sloped gently upward to the center, where a glass insert in the floor allowed visitors to peer down into the entrance tunnel with its laser lights. Other parts of the floor were covered in different materials and textures–stone, wood, carpet. As the visitor moved around, the handset delivered a changing array of sounds. While a viewer stood on the patch of plastic grass, for example, loop antennas embedded in the floor might trigger the sound of birds or a lawn mower.

The experience was deeply personal: You could wander about at your own pace, in any direction, and compose your own trippy sensory experience.

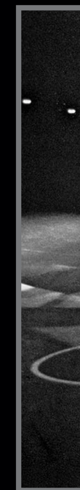
**To pull off such a feat** of techno-art required an extraordinary amount of engineering. The mirror dome alone took months to design and build. E.A.T. viewed the mirror as, in Frosty Myers’s words, the “key to the whole Pavilion,” and it dictated much of what was planned for the interior. The research and testing for the mirror largely fell to members of E.A.T.’s Los Angeles chapter, led by Elsa Garmire. The physicist had done her graduate work at MIT with laser pioneer Charles Townes and then accepted a postdoc in electrical engineering at Caltech. But Garmire found the environment for women at Caltech unsatisfying, and she began to consider the melding of art and engineering as an alternate career path.

After experimenting with different ideas, Garmire and her colleagues designed a mirror modeled after the Mylar balloon satellites launched by NASA. A vacuum would hold the mirror’s Mylar lining in place, while a rigid outer shell held in the vacuum. E.A.T. unveiled a full-scale prototype of the mirror in September 1969 in a hangar at a Marine Corps airbase. It was built by G.T. Schjeldahl Co., the Minnesota-based company responsible for NASA’s Echo and PAGEOS balloon satellites. Gene Youngblood, a columnist for an underground newspaper, found himself mesmerized when he ventured inside the “giant womb-mirror” for the first time. “I’ve never seen anything so spectacular, so transcendentally surrealistic.... The effect is mind-shattering,” he wrote. What you saw depended on the ambient lighting and where you were standing, and so the dome fulfilled E.A.T.’s goal of providing each visitor with a unique, interactive experience. Such effects didn’t come cheap: By the time Expo ’70 started, the cost of the pavilion’s silver lining came to almost $250,000.

An even more visually striking feature of the pavilion was its exterior fog. Ethereal in appearance, it required considerable real-world engineering to execute. This effort was led by Japanese artist Fujiko Nakaya, who had met Klüver in 1966 in New York City, where she was then working. Born in 1933 on the northern island of Hokkaido, she was the daughter of Ukichiro Nakaya, a Japanese physicist famous for his studies of snow crystals. When E.A.T. got the Pepsi commission,

# III. Inside the Experience

A PAVILION OF ONE'S OWN: The goal of the pavilion was to give each of the hundreds of thousands of visitors a personalized, interactive experience. Upon entering, visitors were bathed in patterns of laser light [top, far right]. A giant spherical mirror at the pavilion's center [right] made people appear to float on the ceiling. Visitors carried wireless handsets [bottom, far right], which emitted sounds and noises depending on where you stood in the pavilion.



Klüver asked Fujiko to explore options for enshrouding the pavilion in clouds.

Nakaya's aim was to produce a "dense, bubbling fog," as she wrote in 1972, for a person "to walk in, to feel and smell, and disappear in." She set up meteorological instruments at the pavilion site to collect baseline temperature, wind, and humidity data. She also discussed several ways of generating fog with scientists in Japan. One idea they considered was dry ice. Solid chunks of carbon dioxide mixed with water or steam could indeed make a thick mist. But the expo's health officials ruled out the plan, claiming the massive release of $CO_2$ would attract mosquitoes.

Eventually, Nakaya decided that her fog would be generated out of pure water. For help, she turned to Thomas R. Mee, a physicist in the Pasadena area whom Elsa Garmire knew. Mee had just started his own company to make instruments for weather monitoring. He had never heard of Klüver or E.A.T., but he knew of Nakaya's father's pioneering research on snow.

Mee and Nakaya figured out how to create fog by spraying the water under high pressure through copper lines fitted with very narrow nozzles. The lines hugged the edges of the geodesic structure, and the 2,500 or so nozzles atomized some 41,600 liters of water an hour. The pure white fog spilled over the structure's angled and faceted roof and drifted gently over the fairground. Breer compared it to the clouds found in Edo-period Japanese landscape paintings.

**While the fog** and mirrored dome were the pavilion's most obvious features, hidden away in a control room sat an elaborate computerized sound system.

Designed by Tudor, the system could accept signal inputs from 32 sources, which could be modified, amplified, and toggled among 37 speakers. The sources could be set to one of three modes: "line sound," in which the sound switched rapidly from speaker to speaker in a particular pattern; "point sound," in which the sound emanated from one speaker; and "immersion" or "environmental" mode, where the sound seemed to come from all directions. "The listener would have the impression that the sound was

CLOCKWISE FROM LEFT: FUJIKO NAKAYA/E.A.T./GETTY RESEARCH INSTITUTE, LOS ANGELES, 2014.R.20 (2); SHUNK-KENDER/J. PAUL GETTY TRUST/GETTY RESEARCH INSTITUTE, LOS ANGELES, 2014.R.20

POST YOUR COMMENTS AT
spectrum.ieee.org/pepsipavilion-mar2020

somehow embodied in a vehicle that was flying about him at varying speeds," Tudor explained.

The audio system also served as an experimental lab. Much as researchers might book time on a particle accelerator or a telescope, E.A.T. invited "resident programmers" to apply to spend several weeks in Osaka exploring the pavilion's potential as an artistic instrument. The programmers would have access to a library of several hundred "natural environmental sounds" as well as longer recordings that Tudor and his colleagues had prepared. These included bird calls, whale songs, heartbeats, traffic noises, foghorns, tugboats, and ocean liners. Applicants were encouraged to create "experiences that tend toward the real rather than the philosophical." Perhaps in deference to its patron's conservatism, E.A.T. specified it was "not interested in political or social comment."

**In sharp contrast** to E.A.T.'s sensibilities, Pepsi executives didn't view the pavilion as an experiment or even a work of art but rather as a product they had paid for. Eventually, they decided that they were not well pleased by what E.A.T. had delivered. On 20 April 1970, little more than a month after the pavilion opened to the public, Pepsi informed Klüver that E.A.T.'s services were no longer needed. E.A.T. staff who had remained in Osaka to operate the pavilion smuggled the audio tapes out, leaving Pepsi to play a repetitive and banal soundtrack inside its avant-garde building for the remaining months of the expo.

Despite E.A.T.'s abrupt ouster, many critics responded favorably to the pavilion. A *Newsweek* critic called it "an electronic cathedral in the shape of a geodesic dome," neither "fine art nor engineering but a true synthesis." Another critic christened the pavilion a "total work of art"—a *Gesamtkunstwerk*—in which the aesthetic and technological, human and organic, and mechanical and electric were united.

In hindsight, the Pepsi Pavilion was really the apogee for the art-and-technology movement that burst forth in the mid-1960s. This first wave did not last. Some critics contended that in creating corporate-sponsored large-scale collaborations like the pavilion, artists compromised themselves aesthetically and ethically—"freeload[ing] at the trough of that techno-fascism that had inspired them," as one incensed observer wrote. By the mid-1970s, such expensive and elaborate projects had become as discredited and out of fashion as moon landings.

Nonetheless, for many E.A.T. members, the Pepsi Pavilion left a lasting mark. Elsa Garmire's artistic experimentation with lasers led to her cofounding a company, Laser Images, which built equipment for laser light shows. Riffing on the popularity of planetarium shows, the company named its product the "laserium," which soon became a pop-culture fixture.

Meanwhile, Garmire shifted her professional energies back to science. After leaving Caltech for the University of Southern California, she went on to have an exceptionally successful career in laser physics. She served as engineering dean at Dartmouth College and president of the Optical Society of America. Years later, Garmire said that working with artists influenced her interactions with students, especially when it came to cultivating a sense of play.

After Expo '70 ended, Mee filed for a U.S. patent to cover an "Environmental Control Method and Apparatus" derived from his pavilion work. As his company, Mee Industries, grew, he continued his collaborations with Nakaya. Even after Mee's death in 1998, his company contributed hardware to installations Nakaya designed for the Guggenheim Museum in Bilbao, Spain. More recently, her *Fog Bridge* was integrated into the Exploratorium building in San Francisco.

Billy Klüver insisted that the success of his organization would ultimately be judged by the degree to which it became redundant. By that measure, E.A.T. was indeed a success, even if events didn't unfold quite the way he imagined. At universities in the United States and Europe, dozens of programs now explore the intersections of art, technology, engineering, and design. It's common these days to find tech-infused art in museum collections and adorning public spaces. Events like Burning Man and its many imitators continue to explore the experimental edges of art and technology—and to emphasize the process over the product.

And that may be the legacy of the pavilion and of E.A.T.: They revealed that engineers and artists could forge a common creative culture. Far from being worlds apart, their communities share values of entrepreneurship, adaptability, and above all, the collective desire to make something beautiful. ■

## AUTONOMOUS VEHICLES

CONTINUED FROM PAGE 39

autonomous driving on normal roads and highways.

We and our colleagues at PerceptIn have been trying to address these challenges by focusing on small, slow-speed vehicles that operate in limited areas and don't have to mix with high-speed traffic–university campuses and industrial parks, for example.

The main tactic we've used to reduce costs is to do away with lidar entirely and instead use more affordable sensors: cameras, inertial measurement units, satellite positioning receivers, wheel encoders, radars, and sonars. The data that each of these sensors provides can then be combined though a process called sensor fusion.

With a balance of drawbacks and advantages, these sensors tend to complement one another. When one fails or malfunctions, others can take over to ensure that the system remains reliable. With this sensor-fusion approach, sensor costs could drop eventually to something like $2,000.

Because our vehicle runs at a low speed, it takes at the very most 7 meters to stop, making it much safer than a normal car, which can take tens of meters to stop. And with the low speed, the computing systems have less severe latency requirements than those used in high-speed autonomous vehicles.

PerceptIn's vehicles use satellite positioning for initial localization. While not as accurate as the systems found on highway-capable autonomous cars, these satellite-navigation receivers still provide submeter accuracy. Using a combination of camera images and data from inertial measurement units (in a technique called visual inertial odometry), the vehicle's computer further improves the accuracy, fixing position down to the decimeter level.

For imaging, PerceptIn has integrated four cameras into one hardware module. One pair faces the front of the vehicle, and another pair faces the rear. Each pair of cameras provides binocular vision, allowing it to capture the kind of spatial information normally given by lidar. What's more, the four cameras together can capture a 360-degree view of the environment, with enough overlapping spatial regions between frames to ensure that visual odometry works in any direction.

Even if visual odometry were to fail and satellite-positioning signals were to drop out, all wouldn't be lost. The vehicle could still work out position updates using rotary encoders attached to its wheels—following a general strategy that sailors used for centuries, called dead reckoning.

Data sets from all these sensors are combined to give the vehicle an overall understanding of its environment. Based on this understanding, the vehicle's computer can make the decisions it requires to ensure a smooth and safe trip.

The vehicle also has an anti-collision system that operates independently of its main computer, providing a last line of defense. This uses a combination of millimeter-wave radars and sonars to sense when the vehicle is within 5 meters of objects, in which case it's immediately stopped.

Relying on less expensive sensors is just one strategy that PerceptIn has pursued to reduce costs. Another has been to push computing to the sensors to reduce the demands on the vehicle's main computer, a normal PC with a total cost less than $1,500 and a peak system power of just 400 W.

PerceptIn's camera module, for example, can generate 400 megabytes of image information per second. If all this data were transferred to the main computer for processing, that computer would have to be extremely complex, which would have significant consequences in terms of reliability, power, and cost. PerceptIn instead has each sensor module perform as much computing as possible. This reduces the burden on the main computer and simplifies its design.

More specifically, a GPU is embedded into the camera module to extract features from the raw images. Then, only the extracted features are sent to the main computer, reducing the data-transfer rate a thousandfold.

Another way to limit costs involves the creation and maintenance of the HD maps. Rather than using vehicles outfitted with lidar units to provide map data, PerceptIn enhances existing digital maps with visual information to achieve decimeter-level accuracy.

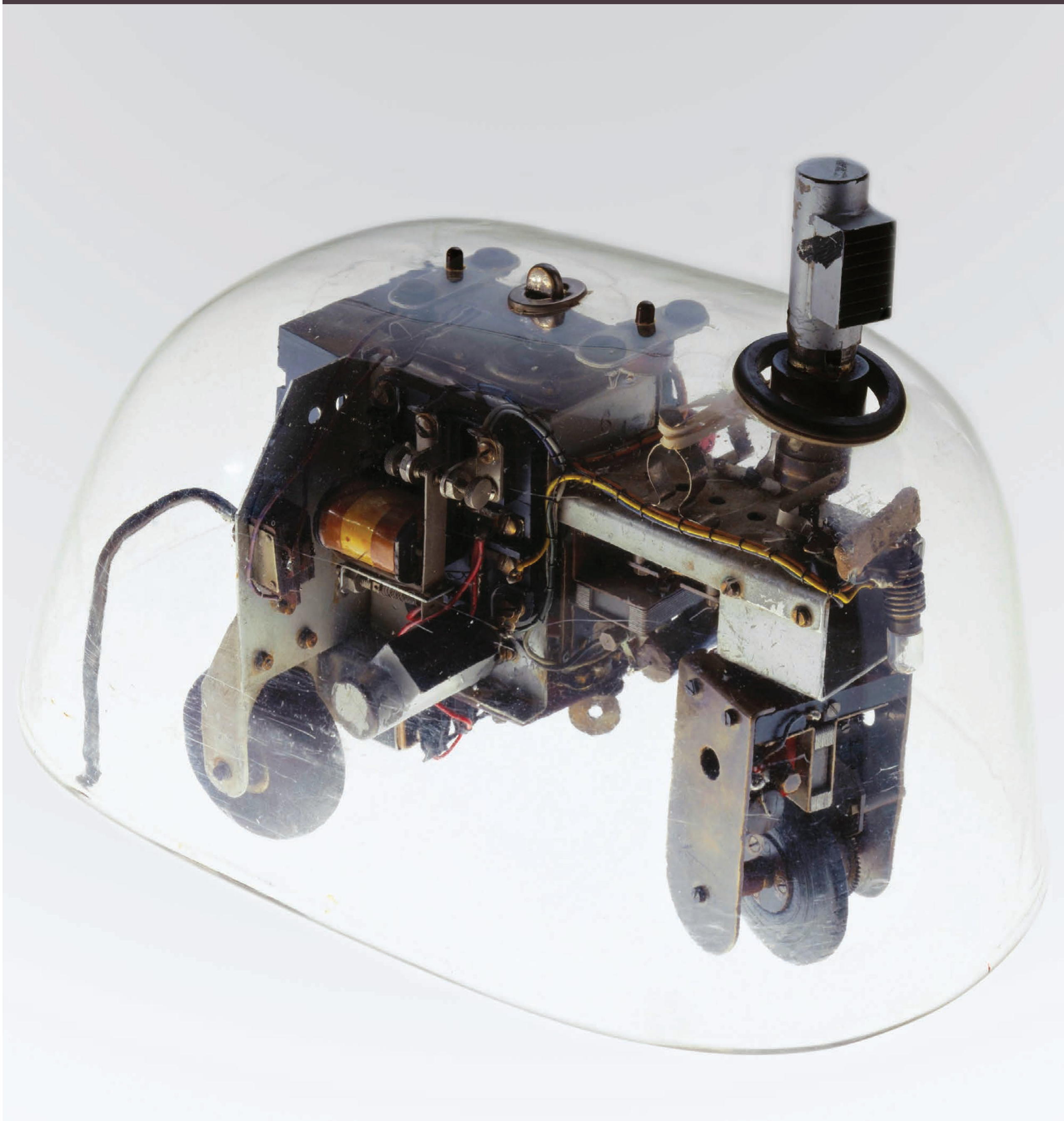
**POST YOUR COMMENTS AT**
spectrum.eee.org/lowspeedautonomy-mar2020

The resultant high-precision visual maps, like the lidar-based HD maps they replace, consist of multiple layers. The bottom layer can be any existing digital map, such as one from the OpenStreetMap project. This bottom layer has a resolution of about 1 meter. The second layer records the visual features of the road surfaces to improve mapping resolution to the decimeter level. The third layer, also saved at decimeter resolution, records the visual features of other parts of the environment—such as signs, buildings, trees, fences, and light poles. The fourth layer is the semantic layer, which contains lane markings, traffic sign labels, and so forth.

**While there's been** much progress over the past decade, it will probably be another decade or more before fully autonomous cars start taking to most roads and highways. In the meantime, a practical approach is to use low-speed autonomous vehicles in restricted settings. Several companies, including Navya, EasyMile, and May Mobility, along with PerceptIn, have been pursuing this strategy intently and are making good progress.

Eventually, as the relevant technology advances, the types of vehicles and deployments can expand, ultimately to include vehicles that can equal or surpass the performance of an expert human driver.

PerceptIn has shown that it's possible to build small, low-speed autonomous vehicles for much less than it costs to make a highway-capable autonomous car. When the vehicles are produced in large quantities, we expect the manufacturing costs to be less than $10,000. Not too far in the future, it might be possible for such clean-energy autonomous shuttles to be carrying passengers in city centers, such as Manhattan's central business district, where the average speed of traffic now is only 7 miles per hour. Such a fleet would significantly reduce the cost to riders, improve traffic conditions, enhance safety, and improve air quality to boot. Tackling autonomous driving on the world's highways can come later. ■



SCIENCE AND SOCIETY PICTURE LIBRARY/GETTY IMAGES

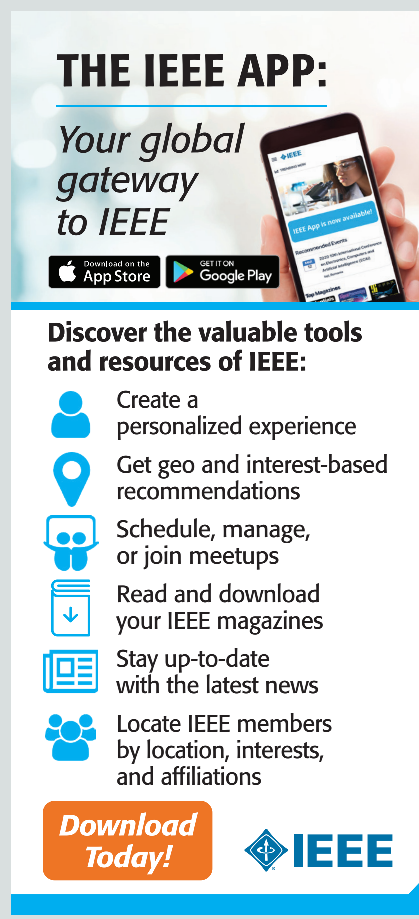
# THE PROTO-ROOMBA

**In the robotics family tree,** Roomba's ancestor was probably the cybernetic tortoise invented in the 1940s by neurophysiologist W. Grey Walter. The turtle-shaped robot could "see" by means of a rotating photocell that steered it toward a light source. If the light was too bright, it would retreat and continue its exploration in a new direction. Likewise, when it ran into an obstacle, a touch sensor would compel the tortoise to reverse and change course. In this way, the robot explored its surroundings. Designed to model simple animal behavior, the tortoise may have satisfied Walter's scientific curiosity, but he never got it to vacuum his floor. ■

↗ **For more on Walter's cybernetic tortoises, see spectrum.ieee.org/pastforward-mar2020**

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal
Best New Journal in STM 2015
Become a published author in 4 to 6 weeks.
Published online only, IEEE Access is ideal for authors who want to quickly announce recent developments, methods, or new products to a global audience.
▪ Submit multidisciplinary articles that do not fit neatly in traditional journals
▪ Reach millions of global users through the IEEE Xplore® digital library with free access to all
Included in Web of Science and has an Impact Factor
Learn more at: ieeeaccess.ieee.org
IEEE


One of the most influential reference resources for engineers around the world.
For over 100 years, Proceedings of the IEEE has been the leading journal for engineers looking for in-depth tutorial, survey, and review coverage of the technical developments that shape our world.
To learn more and start your subscription today, visit
ieee.org/proceedings-subscribe
IEEE

THE IEEE APP:
Your global gateway to IEEE
Download on the App Store
GET IT ON Google Play
Discover the valuable tools and resources of IEEE:
Create a personalized experience
Get geo and interest-based recommendations
Schedule, manage, or join meetups
Read and download your IEEE magazines
Stay up-to-date with the latest news
Locate IEEE members by location, interests, and affiliations
Download Today!
IEEE

MathWorks®

# MATLAB SPEAKS DEEP LEARNING

With MATLAB,® you can build and deploy deep learning models for signal processing, reinforcement learning, automated driving, and other applications. Preprocess data, train models, generate code for GPUs, and deploy to production systems.

mathworks.com/deeplearning