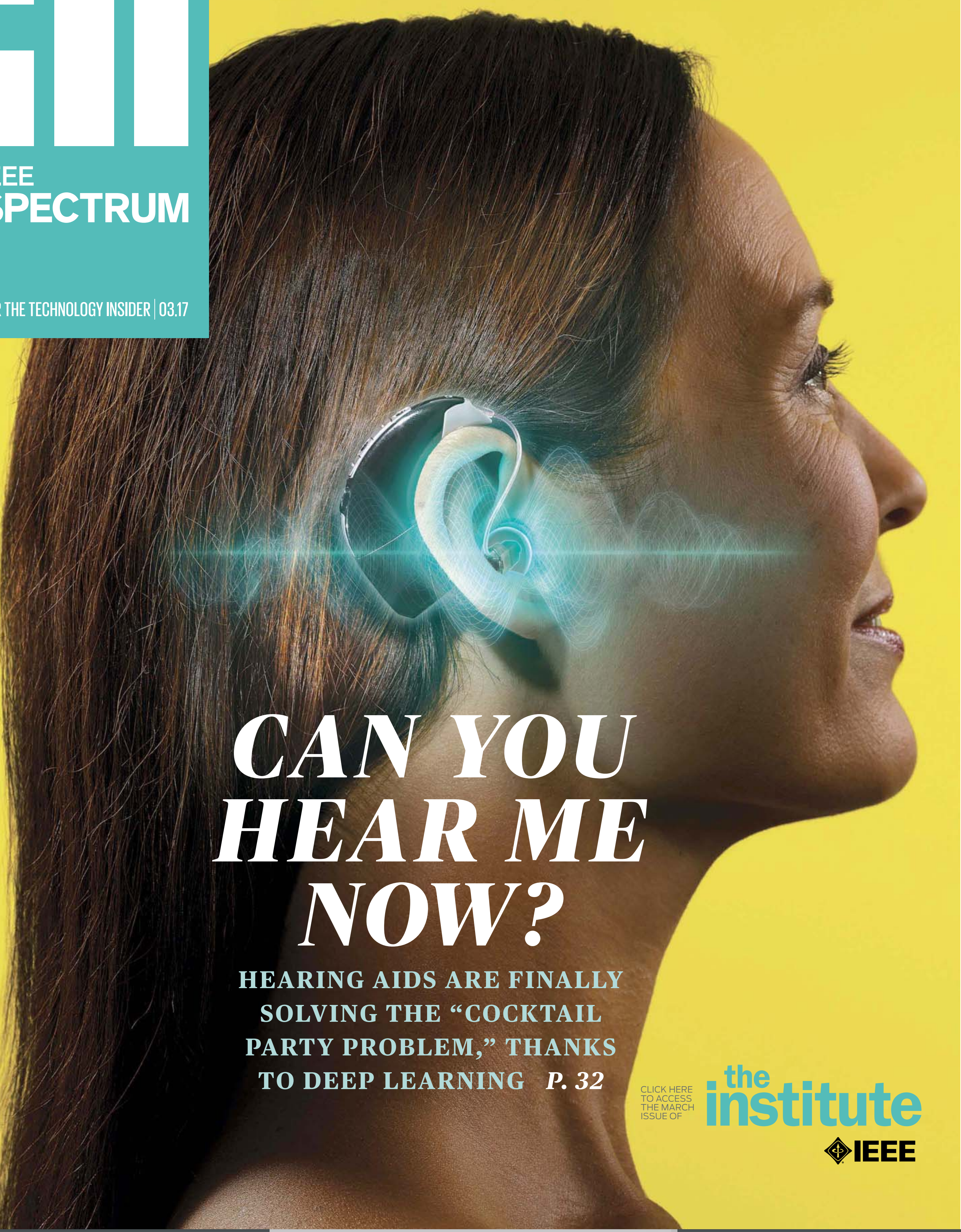
IEEE SPECTRUM

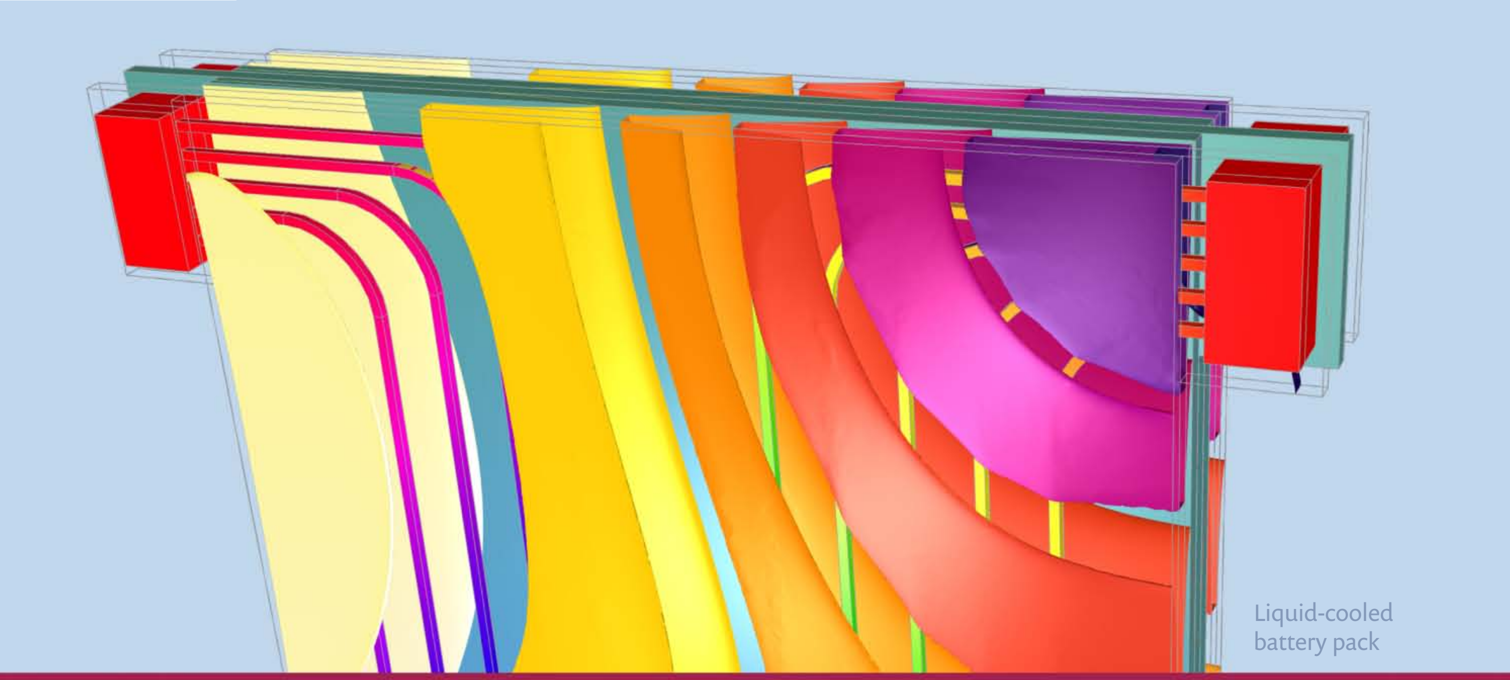
FOR THE TECHNOLOGY INSIDER | 03.17



CAN YOU HEAR ME NOW?

HEARING AIDS ARE FINALLY SOLVING THE "COCKTAIL PARTY PROBLEM," THANKS TO DEEP LEARNING *P. 32*

CLICK HERE TO ACCESS THE MARCH ISSUE OF the institute IEEE



Liquid-cooled battery pack

MULTIPHYSICS FOR EVERYONE

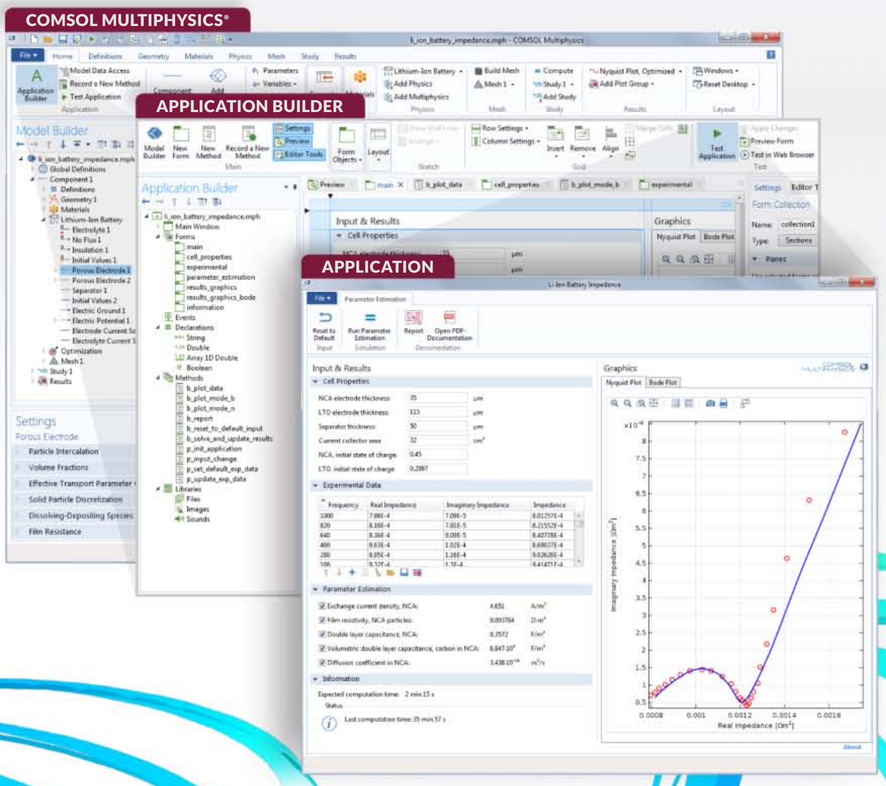
The evolution of computational tools for numerical simulation of physics-based systems has reached a major milestone.

Custom applications are now being developed by simulation specialists using the Application Builder in COMSOL Multiphysics®.

With a local installation of COMSOL Server™, applications can be deployed within an entire organization and accessed worldwide.

Make your organization truly benefit from the power of analysis.

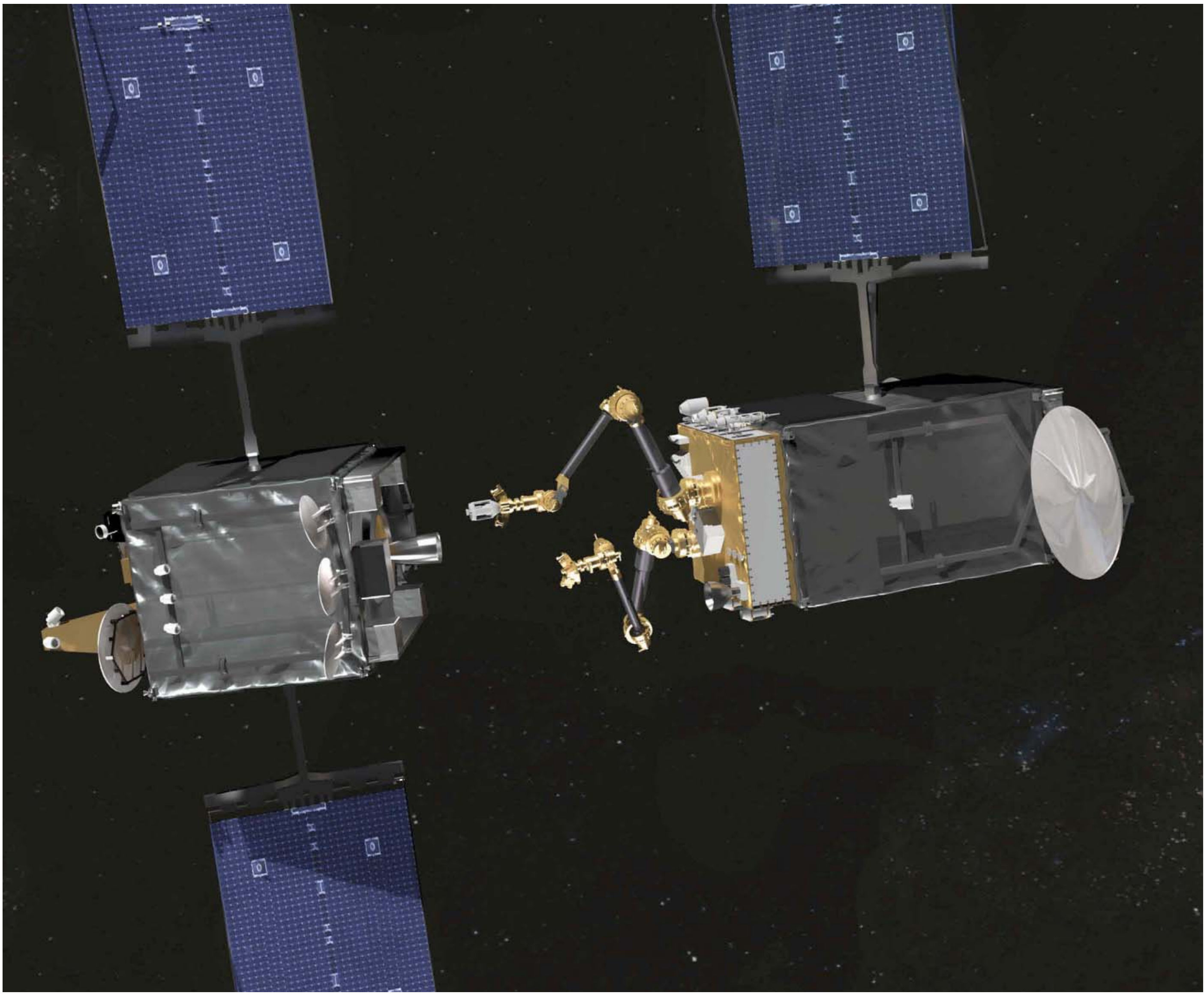
comsol.com/application-builder



Li-ion battery

COMSOL

IEEE SPECTRUM



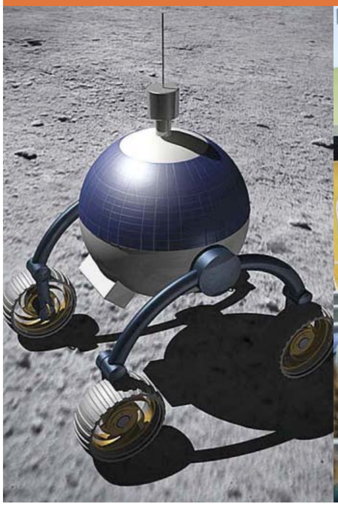
On the cover **Photograph by Dan Saelinger/Trunk Archive**

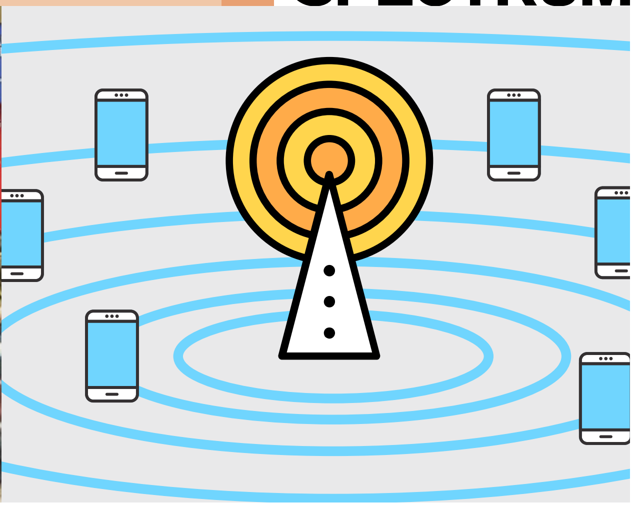
ILLUSTRATION BY **Defense Advanced Research Projects Agency**


Discover
CST STUDIO SUITE 2017

DEPARTMENTS_03.17

IEEE SPECTRUM



FROM LEFT: SYNERGY MOON: CHRISTOPHER LAROSA/UNIVERSITY OF CONNECTICUT:BRANDON PALACIO

07 News

Finalists in a Race to the Moon
Five teams must make it to the moon by the end of the year.
By Rachel Courtland

19 Resources

High Pi-delity
Add-on boards can turn a Raspberry Pi into a hi-fi music center.
By Stephen Cass

06 Opinion

Do We Have to Build Robots That Need Rights?
Then we should build ones that won't mind being owned.
By Susan Hassler

Online Spectrum.ieee.org

Explaining 5G
This video shows how five leading 5G technologies work: massive MIMO, millimeter waves, beam-forming, full duplex, and small cells. Use it as a quick primer before the next party where you might rub shoulders with wireless engineers, or share it to help the public understand these terms: http://spectrum.ieee.org/5g0317

ADDITIONAL RESOURCES

Tech Insider / Webinars

Available at spectrum.ieee.org/webinar

- Getting Ahead With Filter Designer 3D—2 March
- Getting Ahead With Antenna Design—9 March
- SERDES Design and Verification Challenges—On Demand
- From 3-Ring Binders to BOM: Lifetime Products' Step-by-Step Journey to Enterprise PLM—On Demand

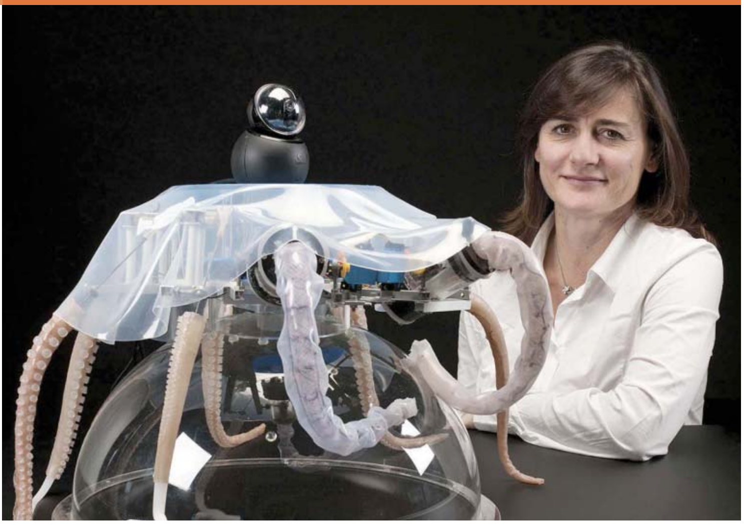
The Institute

Available at theinstitute.ieee.org

- **FUTURE OF NETWORKS** IEEE believes that the fifth generation of wireless networks will be revolutionary, enabling new applications such as the Internet of Things and connected cars. But much work is needed before 5G's full potential can be realized. This special report features IEEE's 5G initiative, an interview with its cofounder, and related conferences and standards.
- **STAY ON TARGET** Entrepreneur and IEEE volunteer Devon Ryan shares tips for how to create a vision for your startup and keep your team motivated to achieve a common goal.
- **HISTORIC BROADCAST** In 1926, members of the Royal Institution of Great Britain witnessed the world's first public demonstration of live television, by inventor John Logie Baird. The event was recently honored with an IEEE Milestone.

IEEE SPECTRUM
(ISSN 0018-9235) is published monthly by The Institute of Electrical and Electronics Engineers, Inc. 3 Park Avenue, New York, NY 10016-5997, U.S.A. Volume No. 54, issue No. 3, International edition. 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 http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

BACK STORY_



JENNIE HILLS/THE LONDON SCIENCE MUSEUM

More Than a Lab Animal

SOMETIMES, TYPICALLY WHILE SHE was chopping up anchovies or adjusting a water tank's pH levels, a thought occurred to Cecilia Laschi: Nothing in her engineering education or professional life had prepared her for duty as an ad hoc aquarist.

Laschi [above], a professor of biorobotics at the Scuola Superiore Sant'Anna in Pisa, Italy, arrived at that unusual juncture through her work on "soft" robotics. She wanted to build supple robots that could bend and move like animals, imagining machines that could maneuver through complicated environments and safely interact with live soft-bodied creatures. In 2006 she decided that the ultraflexible octopus was "the perfect model" for her work and began a long quest to design robotic replicas [see "Octo-Bot," in this issue].

To build these bots, Laschi first needed to know a lot about octopus anatomy. Biology textbooks weren't useful, she discovered, because they didn't include the precise measurements she needed to define the specs for her robots. That's when she moved the first octopus into her lab. Her team conducted experiments to determine how much the animal could elongate and shorten his arms, for example, and how much force he could apply. She called him Alpha and tried to avoid turning him into a pet. "We didn't want to create a relationship with the animals," she says, "because we had to be cold and objective scientists."

But Laschi couldn't help but be impressed by how quickly Alpha learned his experimental tasks, and she sometimes dangled her hand into the tank to say hello to the curious animal. Her team kept Alpha's diet standardized, feeding him uniform chunks of anchovies–but they put the pieces in a screw-top jar to see if Alpha could figure out how to open it. He did. "That was not related to our research," Laschi admits with a laugh. ■

CITING ARTICLES IN IEEE SPECTRUM *IEEE Spectrum* publishes an international and a North American edition, as indicated at the bottom of each page. Both have the same editorial content, but because of differences in advertising, page numbers may differ. In citations, you should include the issue designation. For example, Past Forward is in *IEEE Spectrum*, Vol. 54, no. 3 (INT), March 2017, p. 52, or in *IEEE Spectrum*, Vol. 54, no. 3 (NA), March 2017, p. 56.

CONTRIBUTORS_

Katherine Bourzac

Bourzac, a San Francisco–based journalist, is a veteran at covering the IEEE's International Electron Devices Meeting. One thing she learned at this year's meeting was how easily electrons can break the copper wires used in the slender interconnects in today's advanced chips. "That kind of blew my mind," she says. In "Graphene Could Buttress Next-Gen Computer Chip Wiring" [p. 9], she looks at a potential solution: growing graphene around the copper.

Ben Cutler

Cutler works at Microsoft Research in Redmond, Wash., where he and coauthors Spencer Fowers, Jeffrey Kramer, and Eric Peterson have been exploring the possibility of building data centers underwater [see "Dunking the Data Center," p. 26]. Within 12 months of taking on that daunting assignment, they had a working prototype serving up data from beneath the Pacific. "It was a bit of a Frankenstein thing, but it worked even better than we had expected," says Cutler.

Gordon Roesler

In this issue, Roesler and team members Paul Jaffe and Glen Henshaw write about RSGS, a robotic mission to assist satellites in geosynchronous orbit [p. 44]. Roesler heads RSGS for the U.S. Defense Advanced Research Projects Agency, where he also oversaw earlier robotic-servicing research. "Things have changed a lot," he says. "Ten years ago people were asking, 'Do we really need this?' Today people ask me, 'Can you get on orbit any sooner?' "

Robert Ubell

Ubell is vice dean emeritus of online learning at NYU's Tandon School of Engineering. In this issue, he writes about missteps of early MOOCs (massive open online courses) [p. 22]. Ubell has championed virtual education for nearly two decades, and it's the subject of his recent book of essays, *Going Online: Perspectives on Digital Learning* (Routledge). One of Ubell's first articles on digital education, "Engineers Turn to eLearning," was published in *IEEE Spectrum* in 2000.

DeLiang Wang

Wang is a professor of computer science and engineering at Ohio State University in Columbus, where he directs the Perception and Neurodynamics Laboratory. The lab focuses on algorithms to solve problems related to machine perception. Wang writes in this issue about his work to improve hearing aids [p. 32], having been inspired by his own mother's hearing loss. She struggles to understand speech obscured by background noise, and no hearing aid has helped—yet.

GETTY IMAGES

Do We Have to Build Robots That Need Rights?

Since robots are owned, we should continue to build ones that won't mind being owned

In January, the Legal Affairs Committee of the European Parliament put forward a draft report urging the creation and adoption of EU-wide rules to corral the myriad issues arising from the widespread use of robots and AI–a development, it says, that is "poised to unleash a new industrial revolution." ● It's an interesting read, and a valiant effort to get a handle on how to standardize and regulate the ever-expanding robot universe: drones, industrial robots, care robots, medical robots, entertainment robots, robots in farming–you name it, they're all in there. ● Beginning with Frankenstein's monster, Prague's golem, and Karel Čapek's robot and ending with a code of ethics for robotics engineers and some daunting lists of "shoulds" for robot designers and end users, the 22-page worry catalog toggles between practical concerns about liability, accountability, and safety–who's going to pay when a robot or a self-driving car has an accident?–and far-ranging ones about when robots will need to be designated "electronic persons," and how we will ensure that their creators make them good ones. ● The practical concerns addressed include a call for the creation of a European agency for robotics and artificial intelligence to support the European Commission in its regulation- and legislation-making efforts. Definitions and classifications of robots and smart robots need to be detailed, and a robot registration system described. Interoperability and access to code and intellectual property rights are addressed. Even the impact of robotics on the workforce and the economy are flagged for oversight. ● The "electronic persons" discussion, tucked halfway through the report, caught everyone's attention–perhaps because it's much more fun to catastrophize about HAL 9000 and Skynet than it is to ponder robot insurance requirements. And because personhood–what it legally means to be recognized as a person–is such a loaded topic.

Mady Delvaux, a Luxembourg member of the EP and the report's author, attempted to clarify the designation of what a limited "electronic personality" would be, saying that it would be comparable to the standing that corporations have as legal persons, making it possible for them to conduct business, limit liability, and sue or be sued for damages.

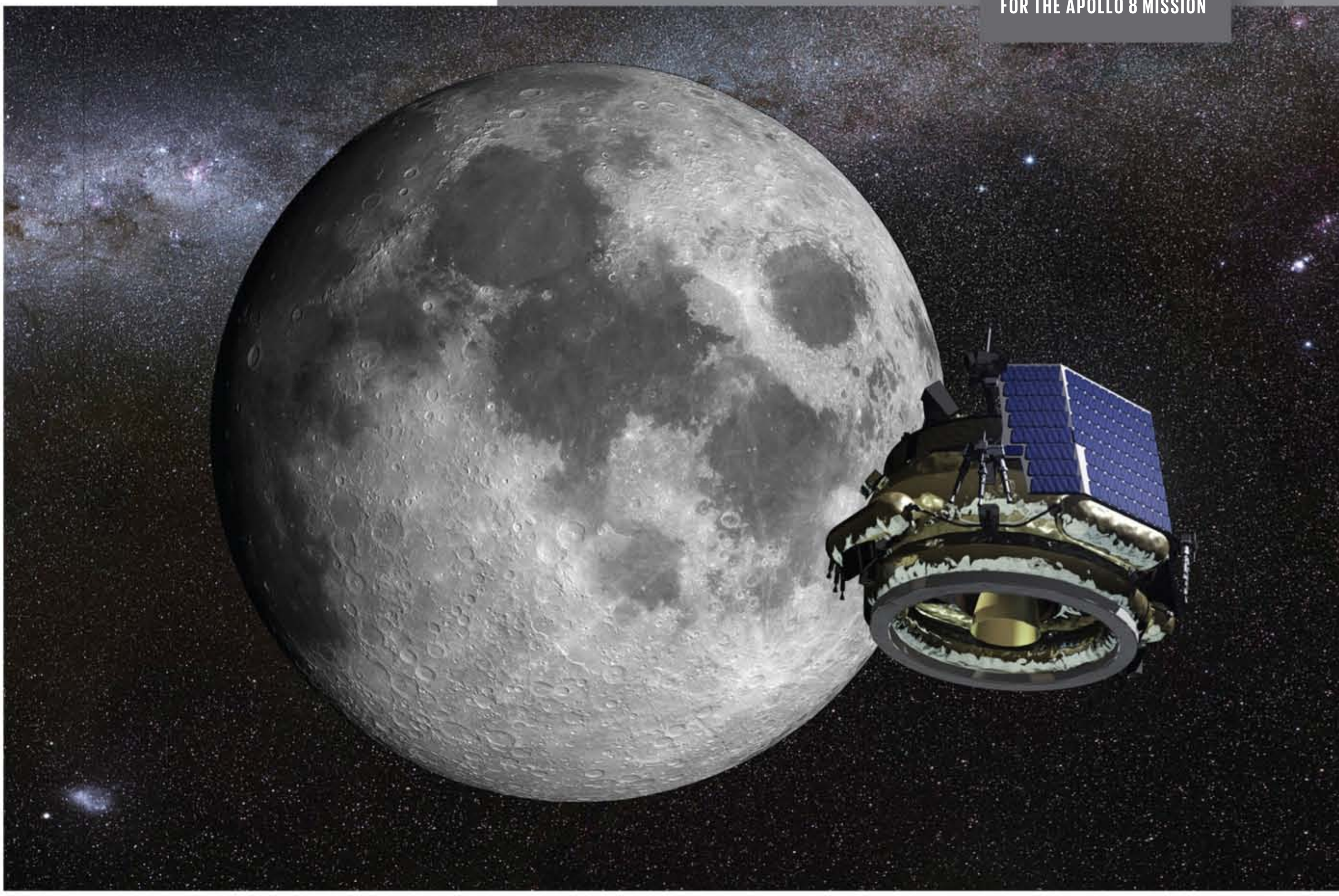
But we haven't finished addressing legal definitions of personhood for women, children, and higher-order animals like chimpanzees yet. Are we really ready to take on robot e-personhood?

I called Joanna Bryson, reader in the department of computer science at the University of Bath, in England, and a working member of the IEEE Ethically Aligned Design project, to ask her what she thought, having just read the Reddit Science "Ask Me Anything" she did about the future of AI and robotics. Her response? As soon as you put the word "person" in the draft, you're probably in trouble.

She told me about Australian law professor S.M. Solaiman's article "Legal Personality of Robots, Corporations, Idols and Chimpanzees: A Quest for Legitimacy," which argues that corporations are legal persons but AIs and chimpanzees aren't. Legal persons must know and be able to claim their rights: They must be able to assert themselves as members of a society, which is why nonhuman animals (and some incapacitated humans), and artifacts like AIs should not, according to Solaiman, be considered legal persons.

But then Bryson said something I had not considered. Since robots are owned–they are in a sense our machine slaves–we can choose *not* to build robots that would mind being owned. We aren't obliged to build robots that we end up feeling obliged to, says Bryson. So instead of assuming that an ethically challenged future saturated with sentient machines is inevitable, we could choose to maintain agency over the machines we are building and defy the technological imperative. Could we do it? Or are we so in thrall to the notion of creating artificial life, monsters, and golems, that it's irresistible?
–SUSAN HASSLER

NEWS



FINISH LINE LOOMS FOR GOOGLE LUNAR XPRIZE

Five teams, four rockets, and 380,000 kilometers to go

MOON EXPRESS

TO THE MOON: A lander by Moon Express is one of five contenders for a US $20 million, 380,000-kilometer race that could end this year.

After nearly 10 years and multiple deadline extensions, the moon seems to be coming into focus for competitors in the Google Lunar XPrize.

Five finalists will have until 31 December to leave the launchpad in pursuit of the US $20 million top prize, which will go to the first team that lands a vehicle on the moon and completes a set of mission requirements, including moving at least 500 meters across the surface and transmitting high-definition video back to Earth.

Expect to see teams gearing up to launch in the last quarter of 2017, says the competition's senior director, Chanda Gonzales-Mowrer: "We're really in the final stretch now."

Each of the finalists has a signed launch contract that has been verified by Google Lunar XPrize. Three have secured spots on well-established launch vehicles. SpaceIL, an Israeli nonprofit, »

Moon Express (United States)

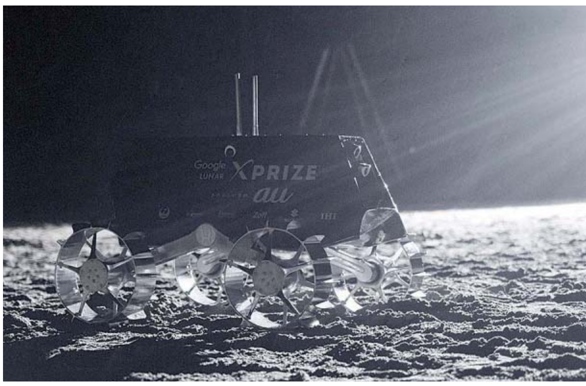
Launch vehicle: Rocket Lab Electron

The XPrize is only part of the plan for Moon Express, which aims to build a business around lunar resources. It plans a series of launches with the new Electron rocket. Moon Express aims to use its lunar lander's thrusters to "hop" to the finish.

TeamIndus (India)

Launch vehicle: Indian Space Research Organization PSLV-XL

Bangalore-based TeamIndus is targeting a launch in late December aboard a workhorse rocket operated by India's space agency. The team's lander will deploy a small rover weighing roughly 5 kilograms.



Hakuto (Japan)

Launch vehicle: Indian Space Research Organization PSLV-XL

After losing its original ride with competitor Astrobotic, which withdrew from the race, Hakuto found a new berth with TeamIndus.

Synergy Moon (International)

Launch vehicle: Interorbital Systems Neptune 8

Synergy Moon plans to use a new launch vehicle to send it directly into a lunar trajectory from Earth's surface, without stopping off in Earth orbit. The team is aiming for another moon mission this year; both will carry one rover.

SpaceIL (Israel)

Launch vehicle: SpaceX Falcon 9

The only nonprofit in the mix, SpaceIL aims to give Israel an "Apollo moment." After landing, the team's spacecraft will use its thrusters to take off again and hop along the surface.

CLOCKWISE FROM TOP LEFT: MOON EXPRESS; TEAMINDUS; SYNERGY MOON; SPACEIL; HAKUTO

will use a SpaceX Falcon 9 to get off the ground. Bangalore-based TeamIndus has secured a berth on a Polar Satellite Launch Vehicle (the PSLV-XL), operated by the Indian Space Research Organization. TeamIndus will carry a rover from Japanese competitor Hakuto with it.

The two other finalists have placed their bets on untried prospects, firms that are aiming for low-cost launches that, as of press time, have yet to reach orbit. The international team Synergy Moon's ride is with Interorbital Systems, a California-based firm developing modular launch vehicles. And Moon Express, a U.S. team building a business in lunar resources development, has penned a three-launch agreement with Rocket Lab, based in the United States and New Zealand. The company has a two-stage launch vehicle called Electron that's designed for small payloads. Rocket Lab says it hopes to start orbital test launches of Electron soon.

The teams are taking different approaches to fulfilling the prize requirements. Some are building rovers to traverse the requisite 500 meters. Others hope to use their lunar landers to cross the distance, by firing thrusters to make a "hop." In the coming year, Gonzales-Mowrer says, teams will work closely with the XPrize organization to determine how judges will affirm that prize requirements have been met.

But the first step is getting off the ground, and it could be a tight race to leave the launchpad by the end of the year. So far, only TeamIndus has announced a specific launch date: 28 December, just three days shy of the deadline. "The teams are under a considerable amount of pressure," Gonzales-Mowrer says, and there are no plans at present to extend the deadline.

The Google Lunar XPrize has already been extended multiple times, most recently from 2015 to 2016 and then again to the end of this year. The extensions have helped teams handle what has turned out to be one of the most difficult aspects of the competition: raising the money. "Funding has been a challenge for teams from the very beginning, and it will continue to be until the end," says Gonzales-Mowrer.

A mission can easily cost more than the top prize. SpaceIL, for example, estimates its mission will cost $70 million. So far the team has raised $50 million of that total, but SpaceIL member Elana Lichtenstein says the team is confident it will be able to make up the difference.

As it can appeal to a higher purpose, SpaceIL may have an easier time than others raising funds for its mission, says Lichtenstein: "We're asking people to make a social and emotional investment in STEM [science, technology, engineering, and math] education for the future of Israel." So far, the team estimates that its educational outreach efforts have reached 250,000 students. It hopes to double that total by the end of this year.

As with other teams, SpaceIL is still readying its spacecraft, which is being built at an Israel Aerospace Industries facility. The team's launch contract is for a six-month window that extends into 2018. If it can't make the 2017 deadline, SpaceIL still aims to go to the moon: "We have an educational mission that we are intent on achieving that has little to do with the time frame of the competition," Lichtenstein says.

Some teams that were expected to advance to the final round did not. The Germany-based group Part Time Scientists, for example, announced last year that it had secured a launch contract. But the team did not meet the prize requirements, because its launch did not have the potential to occur before the 2017 deadline, Gonzales-Mowrer says.

Longtime front-runner U.S.-based Astrobotic Technology is also missing from the list of finalists. The team withdrew from the competition, explaining that the prospect of rushing to make the XPrize deadline conflicted with the company's goal of building a sustainable, long-term business. The group now aims to fly its first mission in 2019. —RACHEL COURTLAND

NEWS

GRAPHENE COULD BUTTRESS NEXT-GEN COMPUTER CHIP WIRING

Current can literally blow copper interconnects away, but graphene could keep them intact

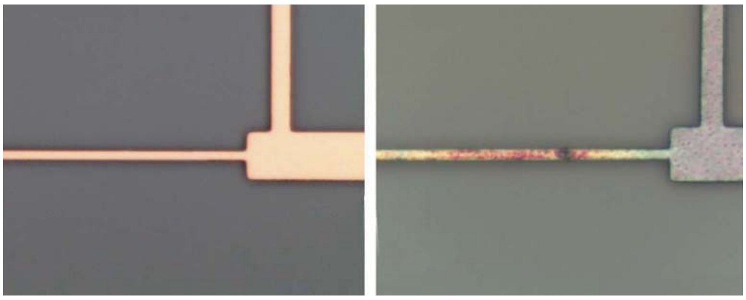
Most of the hand-wringing over the fate of Moore's Law focuses on the ever-shrinking silicon transistor. But increasingly researchers are concerned with another critical part of the infrastructure: the copper wires that connect individual transistors to form complex circuits.

At the IEEE International Electron Devices Meeting in San Francisco in December, researchers described the coming problems for copper interconnects and debated ways of getting around them. One approach, studied by a group led by Stanford electrical engineer H.-S. Philip Wong, is to gird copper with graphene. Wong's group found that the nanomaterial can alleviate a major problem facing copper, called electromigration.

Copper wires are getting so thin, and must carry so much current, that the atoms in the wire can literally get blown out of place. "The electron wind can physically move the copper atoms and create a void," says Wong. Growing graphene around copper wires prevents this, according to research that Wong's group presented at the meeting. It also seems to bring down the resistance of the copper wires.

Ruth Brain, an Intel Fellow and director of interconnect technology and integration at the company's Hillsboro, Ore., location, explained how interconnects are being pushed to their limits. More transistors per chip area means more interconnects have been installed to connect them. The first chips to use copper interconnects, which were produced in 2000, had 1 kilometer of wiring per square centimeter. Today's 14-nanometer-node processors contain more than 10 km of copper wiring in the same area, she said.

In order to improve performance, the ever-narrowing copper wires must carry ever more current. The amount



DEATH BY ELECTROMIGRATION: Copper interconnects are now so narrow that current can cause a break by knocking atoms out of place.

STANFORD UNIVERSITY

of current per area in a wire is called its current density, which has been increasing on cutting-edge chips for two reasons: Wires are shrinking, and higher currents must course through them to enable the faster switching speeds that boost performance.

And herein lies the challenge: The narrower the wire, the higher its resistance. "Interconnects have had to shrink while increasing the current densities by 20 times," said Brain. "You would burn your house down if you did this in your house."

Today's solution is to deposit copper interconnects within trenches lined with 2-nanometer-thick walls of tantalum nitride. This lining keeps the copper from escaping, and Wong says copper will probably endure through the coming 10- and 7-nm nodes. As device features keep shrinking, though, 2-nm walls will be far too thick, says Wong. Researchers are investigating other linings that may prevent electromigration, including ruthenium and magnesium, but at 0.3 nm, he says, graphene is thinner than anything else.

The semiconductor industry avoids integrating new materials as long as possible, but Wong says there isn't much choice in this situation: If copper's life can't be extended, it will have to be replaced with a new material anyway, such as cobalt.

The Stanford group worked with Lam Research Corp., which makes chip manufacturing tools, as well as researchers from Zhejiang University, in China, to make and test the composite interconnects. The materials make a good pair; graphene is often made by growing it on copper. Lam Research has developed a proprietary process for doing this at temperatures that won't damage the rest of the chip—below 400 °C. Compared to copper alone, the composite limited electromigration by a factor of 10. And the composite wires had half the electrical resistance.

Wong says the interconnect problem can no longer be dismissed. "Before, most of the time we were hearing about transistors," he says. "Now it's not just transistors but wires, memory—many other things that were previously not a problem are beginning to be a problem."

—KATHERINE BOURZAC

"Interconnects have had to shrink while increasing the current densities by 20 times. You would burn your house down if you did this in your house"
—Ruth Brain, Intel

AT-HOME ELECTRIC HEADBAND FOR DEPRESSION COULD GO MAINSTREAM

A consumer-friendly gadget could help tDCS treatment catch on

A doctor's prescription for clinical depression could one day sound like this: In the comfort of your own home, slip on a brain-zapping headband a few times per week. For 20 minutes, it will send a tiny stream of electricity through your brain.

The treatment would be delivered by a user-friendly type of brain stimulation called tDCS (transcranial direct-current stimulation). This mind-altering technique has become a hot topic in neuroscience research over the last decade, and it's now beginning the transition from lab to doctor's office.

In South Korea, the startup Ybrain is betting that its tDCS headband, specifically designed to treat depression, will be the product that brings the electric treatment into the medical mainstream. Ybrain founder and CEO Kiwon Lee expected his device to receive regulatory approval in Korea in February. He plans to roll out the device in 70 Korean hospitals this year to reach thousands of patients with clinical depression. The company will use data from all those patients to build a case for approval first in Europe, Lee says, and then in the United States, where the regulatory requirements are most stringent. "After one device is approved [in the United States], it will be seen as a mainstream treatment," he says.

tDCS is considered an exciting new possibility for clinical use because the gear is cheap, portable, and easy to use. The headsets press electrodes against particular locations on the scalp to channel a few milliamperes of current through a specific brain region, and they can be powered by a 9-volt battery. Last year there were nearly 700 papers published about tDCS, with studies on every topic imaginable: Researchers have experimented with physical rehab for stroke patients, memory improvement for people with Alzheimer's disease, and craving reduction for smokers and alcoholics.

Ybrain may not be the first to enter the all-important American market; New York City–based Soterix Medical may beat it. Soterix chief technology officer Abhishek Datta says that its tDCS system for depression was approved for clinical use in Europe in late 2015, and he says more than 30 European clinics now use it. The company is currently pursuing approval in Canada and the United States. At the moment, the Soterix treatment is not quite as user friendly as Ybrain's. Although the company offers a take-home version of its tDCS system, Datta says most European patients are receiving their stimulation at a clinic under a doctor's supervision.

Ybrain hopes its slick, consumer-friendly design will make at-home treatment the norm. "It's designed for home use," says Lee, "so physicians can electronically prescribe the device and patients can bring it to their homes."

To make the medical gadget easy to use and fail-safe, Ybrain's designers kept it simple. "The device has just a power button, no interface," Lee says. "The patient turns it on and puts the device on their head,

YBRAIN

and it automatically starts the stimulation based on the doctor's prescription." If the doctor prescribed one 20-minute session per day at 2 milliamperes, for example, the gadget wouldn't allow more stimulation than the prescribed limit.

Though companies are pushing ahead with commercialization, many scientific questions must still be answered about tDCS. Neuromodulation experts agree that it can alter the brain in meaningful ways—but they're still figuring out which medical conditions will truly benefit from these alterations. An article recently published in the journal *Clinical Neurophysiology* reviewed the evidence on tDCS effectiveness across a range of disorders; for most, the article said that there haven't yet been enough rigorous trials to draw conclusions.

However, that article singled out tDCS treatment for depression as "probably effective." Considering the World Health Organization's estimate that some 350 million people around the world suffer from depression, many researchers and clinicians think that depression might be the first killer app for tDCS.

Andre Brunoni, a neuroscientist and psychiatrist at the University of São Paulo, provided strong evidence of tDCS's utility for depression in a 2013 study he coauthored that compared tDCS to a standard antidepressant medication. That study found electric and pharmaceutical therapies to be equally effective, and it found that patients who received both treatments together fared best of all. But Brunoni says he's not ready to forge ahead with tDCS in his clinical practice; he's still waiting for answers to many crucial questions that will define its practical use. "We don't know the best 'dose' to use for tDCS," he says. "What's the best amount of current to use? What's the optimal number of sessions? There are a lot of parameters to explore."

If such issues are resolved, Brunoni says tDCS may offer several advantages over both antidepressant drugs and another device-based depression treatment, transcranial magnetic stimulation (TMS), which is performed using a large and comparatively expensive machine at a clinic. Because tDCS treatment is relatively cheap, Brunoni says, private medical plans would likely prefer it. Safe and reliable devices from companies such as Soterix and Ybrain will allow home use "with almost zero risk," he says. And treatments can be short-term, beginning with an intense phase—for example, two weeks of daily sessions—followed by less frequent "maintenance" stimulation sessions. Brunoni says this timing offers a big advantage over antidepressant drugs, which require treatment for at least nine months and can cause unpleasant side effects throughout that period (like weight gain and sexual problems).

While hopes are high for tDCS depression treatments, some experts are cautious. Mark George, an eminent neurologist and psychiatrist who directs the Brain Stimulation Laboratory at the Medical University of South Carolina, in Charleston, says he's not yet ready to prescribe tDCS to his clinically depressed patients. He invokes Hippocrates: "First, do no harm." While there's ample evidence that tDCS is safe when used correctly, George notes that physical safety isn't the only consideration. The depressed patients he sees are desperate, despondent, and looking to him for salvation.

"Hope is fragile," George says, "and you want to be very careful with what you throw out there for your patients. You want to go with your best shot. And right now, I don't see tDCS as your best shot." —ELIZA STRICKLAND

BAND OF AID: Ybrain's transcranial direct-current stimulation headband [left] is designed for home use. A doctor programs prescriptions, including amperage and duration of treatment, into the device's processing modules [above].

NEWS

SELF-DRIVING CARS HAVE A BICYCLE PROBLEM

Bikes are hard to spot and hard to predict

ISTOCKPHOTO

Robotic cars are great at monitoring other cars, and they're getting better at noticing pedestrians, squirrels, and birds. The main challenge, though, is posed by the lightest, quietest, swerviest vehicles on the road.

"Bicycles are probably the most difficult detection problem that autonomous vehicle systems face," says Steven Shladover, a research engineer at the University of California, Berkeley.

Nuno Vasconcelos, a visual computing expert at the University of California, San Diego, says bikes are difficult to detect because they are small, fast, and heterogeneous: "A car is basically a big block of stuff. A bicycle has much less mass.... There are more shapes and colors, and people hang stuff on them."

That's why the detection rate for cars has continually outstripped that for bicycles. Most of the improvement has come from techniques whereby systems train themselves by studying thousands of images in which known objects are labeled.

Consider the Deep3DBox algorithm presented recently by researchers at George Mason University and stealth-mode robotic-taxi developer Zoox, based in Menlo Park, Calif. On an industry-recognized benchmark vision test, Deep3DBox identifies 89 percent of cars. Scores below 70 percent prevailed just a few years ago.

Deep3DBox further excels at a tougher task–predicting which way vehicles are facing. It does so by inferring a 3D box around each object spotted on a 2D image. "Deep learning is typically used for just detecting pixel patterns. We figured out an effective way to use the same techniques to estimate geometrical quantities," explains Deep3DBox contributor Jana Košecká, a computer scientist at George Mason University, in Fairfax, Va.

However, when spotting and orienting bikes and bicyclists, the performance plummets. Deep3DBox is among the best, yet it spots only 74 percent of bikes in the

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

benchmark test. And although it can orient over 88 percent of cars, it scores just 59 percent for bikes.

Košecká says commercial systems are delivering better results as developers gather massive proprietary data sets of road images with which to train their systems. And she says most robocars augment their visual processing with lidar and radar, which help recognize bikes and their relative position even if they can't help determine their orientation.

Further strides, meanwhile, are coming via high-definition maps like Israel-based Mobileye's Road Experience Management system. These maps offer computer vision algorithms a head start in identifying bikes, which stand out as anomalies from the prerecorded street views. Ford Motor Co. says "highly detailed 3D maps" are at the core of the 70 self-driving test cars it plans to have on roads this year.

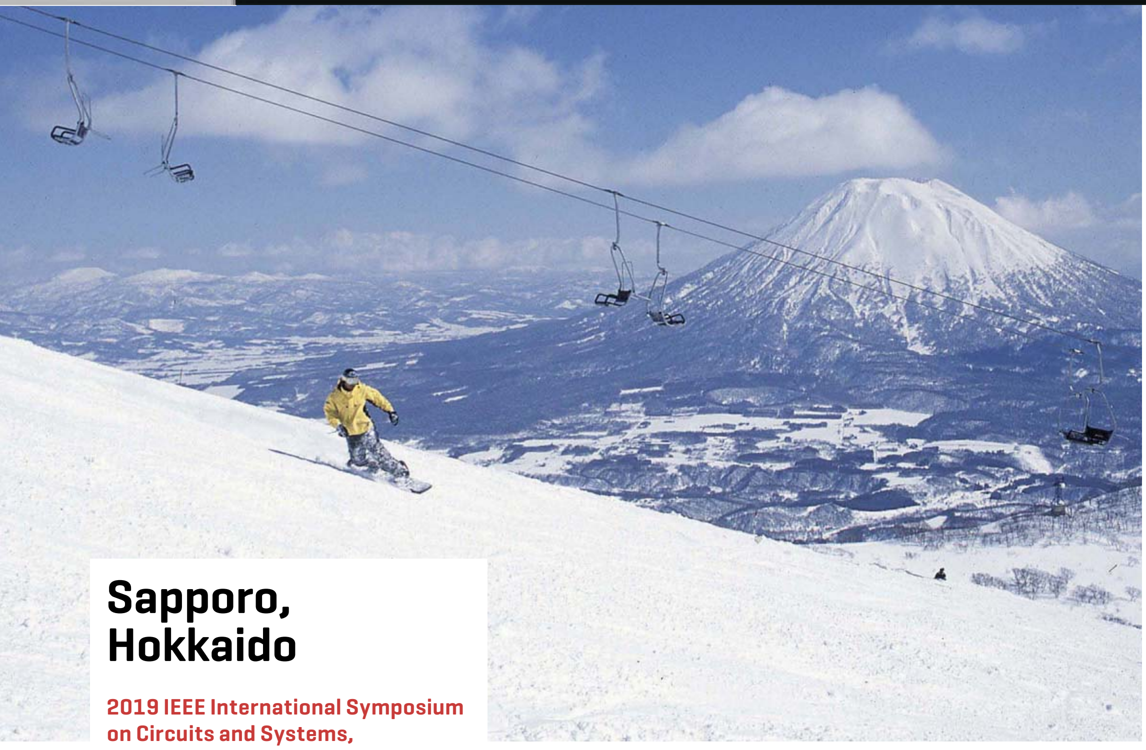
Combine these elements and you can get some impressive results, such as the bike spotting demonstrated last year by Google's vehicles. Waymo, Google's autonomous vehicle spin-off, unveiled proprietary sensors with upgraded bike-recognition capabilities at the Detroit Auto Show, in January.

UC San Diego's Vasconcelos doubts that today's sensing and automation technology is good enough to replace human drivers, but he believes it can already help them avoid accidents. The first commercial applications of cyclist detection are appearing in automated emergency braking (AEB) systems for conventional vehicles. Volvo began offering the first cyclist-aware AEB in 2013; it is rolling out similar tech for European buses this year.

That said, AEB systems suffer from a severe limitation that autonomous vehicle developers are struggling with—predicting where the stuff that's moving around the vehicles will go next. "Bicycles are much less predictable than cars because it's easier for them to make sudden turns or jump out of nowhere," says Košecká.

That means cyclists will be living a while longer with the human error that contributes to 94 percent of traffic crashes, according to U.S. regulators. "Everybody who bikes is excited about the promise of eliminating that," says Brian Wiedenmeier, executive director of the San Francisco Bicycle Coalition. But he says it is right to wait for the technology to mature.

In December, Wiedenmeier warned that self-driving taxis deployed by Uber Technologies were violating laws designed to protect cyclists. He applauded when California officials pulled the vehicles' registrations, citing the firm's refusal to secure state permits for them. "Like any new technology, this needs to be tested very carefully," he says. —PETER FAIRLEY



Sapporo, Hokkaido

2019 IEEE International Symposium on Circuits and Systems,
Sapporo Convention Center
http://www.iscas2019.org/

Like northern New England in the U.S. or the Grenoble region of France (which are on about the same latitude), the island of Hokkaido offers cool forests in summer, snow sports in winter, and blazing fall foliage.

"Visitors of all nations have always selected Sapporo—with its distinctive northern culture, delicious food, and warm hospitality—as the most attractive city in Japan," says Professor Yoshikazu Miyanaga, dean of the Hokkaido University Graduate School of Information Science and Technology, and chair of the organizing committee for the upcoming 2019 IEEE International Symposium on Circuits and Systems (ISCAS 2019). The flagship IEEE Circuits and Systems Society (CASS) conference should attract around 1,500 attendees. "Many IEEE members and others who do not belong to CASS attend to keep up with cutting-edge research."

The Sapporo Convention Center, says Professor Miyanaga, is particularly well suited to international technical symposia.

Getting there is easy, with more than 50 scheduled flights from Tokyo's Haneda Airport to Sapporo's Shin-Chitose International. The center of Sapporo is only 30 to 60 minutes away by rail, bus, or car, and "newcomers can find their way around easily, even without their smartphone GPS apps."

Meetings in Japan:

It's easy to forget how vast Japan is. The archipelago sweeps 2,850 km (a 1,770-mile sweep) from Wakkanai in northern Hokkaido, through Honshu, Kyushu, Shikoku, and Okinawa to the islands of Taketomi and Ishigaki far to the south. Along that arc lies a tremendous variety of landscapes, climates, local cultures, and attractive meeting venues.

In 2016 alone, IEEE societies sponsored 50 international and regional meetings in Japan. Most convened in the capital region, from Tokyo to Kyoto, but tech meeting organizers are increasingly looking to other venues, from the semi-Alpine north, where Hokkaido has hosted 7 major IEEE meetings since 2013, to semitropical Okinawa Prefecture, which has welcomed 10 IEEE conferences in four years. (For a list of IEEE meetings in Japan 2016–2019, see http://bit.ly/2hKoawk.)

"We would like to let organizers know the charms of not only the metropolis but also regional cities in Japan," says Etsuko Kawasaki, Executive Director of the Japan Convention Bureau, part of the Japan National Tourism Organization (JNTO).

Competitive Costs

Once considered an expensive destination, Japan is increasingly competitive, says Ms. Kawasaki. The December 2016 exchange rate was 115 yen per U.S. dollar, compared to 78 yen five years ago. On average, a trip to Tokyo can be significantly less expensive than a visit to Paris, London, or (most expensive) New York City.

And there is, moreover, "a wide range of venues...and you can select a suitable one that meets your budget," particularly outside of the big cities—one more reason to consider Japan's smaller cities and outlying islands.

Getting Around Is Easy

Foreigners are often daunted by the thought of getting around in Japan. They don't need to be.

"Traveling to regional cities is easier than meeting planners and

Ishigaki, Okinawa Prefecture

2014 IEEE Asia Pacific Conference on Circuits and Systems,
ANA InterContinental Ishigaki Resort,
17–20 November 2014

Ishigaki is one of the southernmost islands of Okinawa Prefecture. It's a tropical island of 48,000 people just 200 km (125 mi.) east of Taiwan. To put that in perspective, it is on the same latitude as Key West, Florida, or Morocco's West Sahara province.

"Okinawa in general, and Ishigaki in particular, is one of the best sight-seeing spots in Japan.

It is subtropical, with beautiful coral beaches, and has the facilities and quality services to support a first-class meeting," says Professor Hiroshi Ochi of the Kyushu Institute of Technology, who served as General Chair of the 2014 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS 2014, http://www.apccas2014.org/), held there.

The conference theme was, appropriately, "Smart and Green Circuits and Systems." It drew 310 attendees, who were treated not only to rigorous technical sessions but also a banquet that featured a local traditional dance performance and a welcoming talk by the mayor. "We formed a strong connection with the Okinawa Convention and Visitors Bureau that we expect to be useful in planning future conferences."

There were no complaints about travel or expense, either, says Professor Ochi. "There are many direct flights from most major mainland airports, and low-cost carriers were available."

Ski Slopes to Tropical Beaches

attendees imagine, thanks to highly developed transportation networks throughout the nation," says Ms. Kawasaki. "Attendees can arrive in Hokkaido or Kyushu within 2 hours from the major international airports in Tokyo or Osaka."

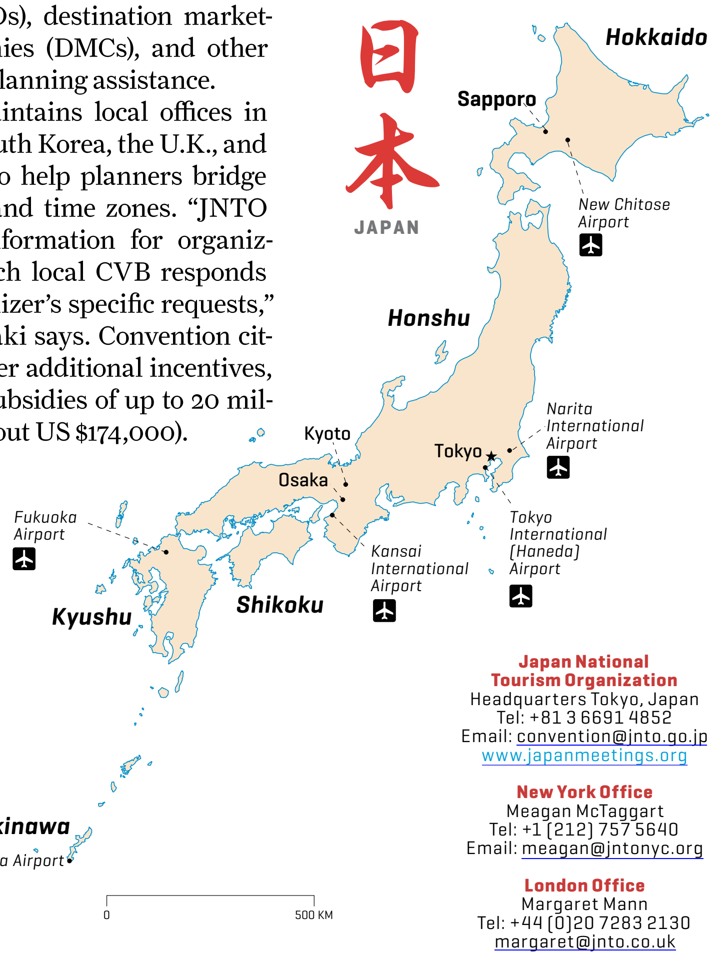
Japan's public transportation system has adopted a common pre-paid integrated-circuit fare card system that can be used to travel on trains, subways, buses, and taxis throughout the country.

The multilingual environment in public places has been improving as Japan prepares for the 2020 Tokyo Olympic and Paralympic Games. Stations and city streets carry multilingual signs, and more tourist information centers open each month.

How JNTO Helps

The Japan Meetings web site (japanmeetings.org) includes tools for searching destinations by facility size, region, and amenities. There are also tools for finding local convention and visitors bureaus (CVBs), professional congress organizers (PCOs), destination marketing companies (DMCs), and other sources of planning assistance.

JNTO maintains local offices in the U.S., South Korea, the U.K., and Singapore to help planners bridge languages and time zones. "JNTO provides information for organizers, and each local CVB responds to the organizer's specific requests," Ms. Kawasaki says. Convention cities often offer additional incentives, including subsidies of up to 20 million yen (about US $174,000).



Japan National Tourism Organization
Headquarters Tokyo, Japan
Tel: +81 3 6691 4852
Email: convention@jnto.go.jp
www.japanmeetings.org

New York Office
Meagan McTaggart
Tel: +1 (212) 757 5640
Email: meagan@jntonyc.org

London Office
Margaret Mann
Tel: +44 (0)20 7283 2130
margaret@jnto.co.uk

PHOTOGRAPH BY Francois Nel/Getty Images

DROID DRIVERS

WHAT YOU'RE LOOKING at is a success story in the ongoing effort to end forced child labor and human trafficking. The robot jockeys shown whipping these camels down the track during the Liwa Sports Festival in the United Arab Emirates in January are direct replacements for the legions of small (and therefore lightweight) boys—many as young as 4 years old—who would otherwise be conscripted from poor nations to egg the beasts on. When countries in the Arabian Peninsula recently began banning this form of abuse, the camel owners turned to technology. An electronic jockey's whipping action is produced by a cordless drill tucked inside its body. The camel's owner controls the electro-jockey from trackside via a repurposed remote car-door opener while racing alongside in an SUV, so that the remote stays within its transmission range.

THE BIG PICTURE NEWS

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

RESOURCES

28.9 kHz: THE MAXIMUM SAMPLING FREQUENCY OF THE AMIGA'S 1985 PAULA, THE FIRST HOME DIGITAL AUDIO CHIP

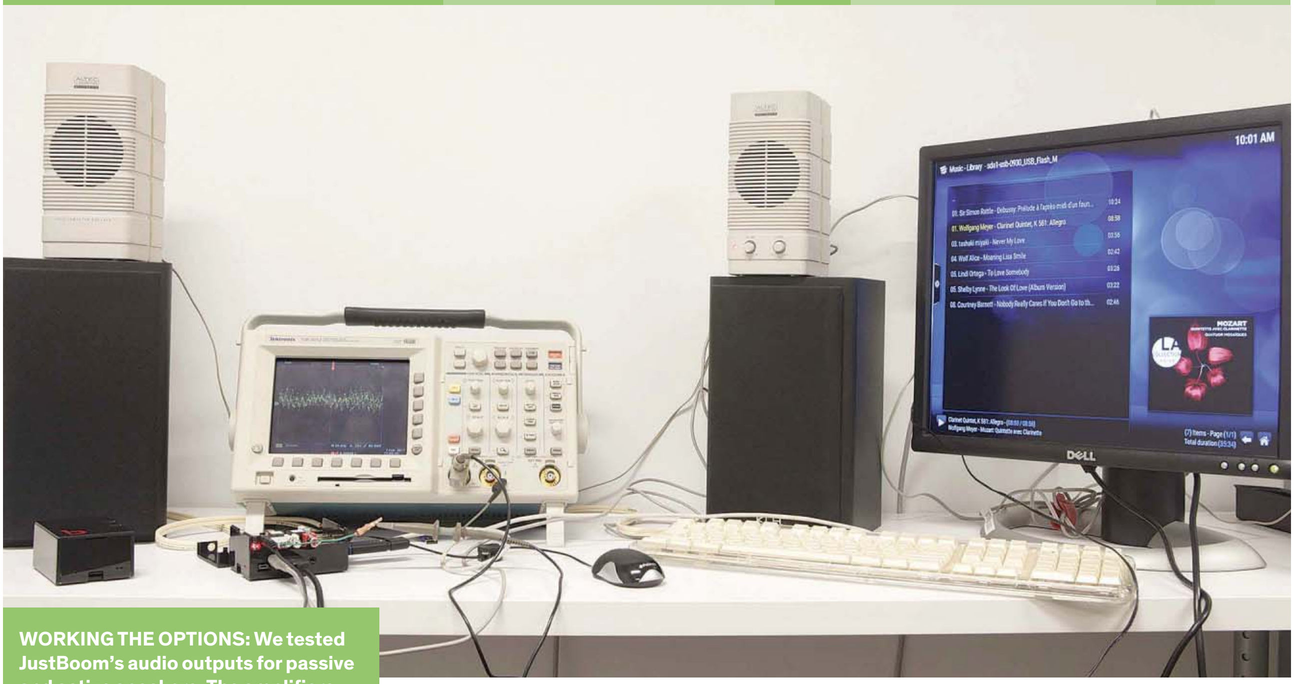
RANDI KLETT

RESOURCES_HANDS ON

HI-FI WITH YOUR PI

JUSTBOOM'S ADD-ON AMPS AND DACS PROVIDE RICH SOUND

From its very first version, the Raspberry Pi attracted interest as a small, inexpensive, and network-friendly home multimedia player. It comes with built-in HDMI and composite video for sending signals to television sets, and most versions have an analog 3.5-millimeter audio socket. Consequently, there's a fleet of Raspberry Pi-compatible software packages designed to play video and audio files using a stripped-down, TV-friendly interface and which can be networked for multiroom audio. ● However, while the high-definition video output is sufficient for viewing most movies (unless you have one of the latest 4K screens, of course), the analog audio is not up to the standards of dedicated audiophiles. As a result, a number of companies have started offering various expansion options to improve the Pi's audio capabilities and ease of integration into existing hi-fi setups. ● We wanted to try out one of these systems at *IEEE Spectrum*—Executive Editor Glenn Zorpette is our resident audiophile—and JustBoom's product line caught my eye due to the flexibility and audio quality promised by its wide range of add-on HAT ("hardware attached on top") and stand-alone boards. JustBoom kindly sent us a selection of its kits, and so we tried them out to see if they could deliver. ● The first thing I realized is that the very flexibility and range that attracted me to JustBoom's products means that figuring out what is supposed to go with what can induce some head-scratching. JustBoom does provide a flowchart-style product guide, but it's somewhat confusing. ● So to simplify matters for Pi owners: If you are looking to connect headphones, active speakers (the kind that have their own power supply), or other audio equipment by way of a 3.5-mm or RCA ▸



WORKING THE OPTIONS: We tested JustBoom's audio outputs for passive and active speakers. The amplifiers proved powerful enough to overload our passive speakers at the highest volumes, so we used an oscilloscope to check for waveform clipping.

RANDI KLETT

socket, you want a DAC HAT, which is a high-audio-resolution digital-to-analog converter. JustBoom's DAC HATs come in two flavors: a US $38 one for all the Model A and Model B versions of the Raspberry Pi, and a $25 one for the smaller Pi Zero. If you want to hook your Pi up to a sound system using an optical audio port, use the Digi HAT, which again comes in two flavors. If you want to power passive speakers (the kind with just two wires running into the back of each) directly from the Pi, go with the $76 Amp HAT, which combines a DAC and an amplifier. (If you want the option of being able to listen using headphones and passive speakers with the same Pi, you can buy a DAC HAT and combine it with a $75 stand-alone Amp board that sits on top of the HAT.) JustBoom sells cases that enclose all the various combinations.

Putting the pieces together was straightforward. Each HAT came with mounting screws and plastic separators to hold the board in place once it was pushed onto the Pi's general purpose input/output (GPIO) connector. Using three Pi's, I assembled a DAC With Amp, an Amp HAT, and a Digi HAT. (I would have tried the supplied HATs for the Pi Zero as well, but Pi Zeros tend to be constantly out of stock!) The amplifier boards have a socket for an external 8- to 24-volt power supply. This also feeds into the Pi, so you can dispense with the USB cable normally used to power the Pi. If you want to deliver the maximum power the amplifiers can output—55 watts peak—you'll need a 24-volt supply rated at over 3.1 amperes.

The cases, each with a nice glossy black finish, snap rather than screw together. (I personally prefer screws, as you avoid having to hold pieces together in just the right way and with just the right amount of pressure lest they pop out at the wrong moment during assembly.)

Installing the software was also straightforward: I formatted a microSD card and copied over a NOOBS boot image from the Raspberry Pi Foundation website. NOOBS lets a user easily install a number of operating systems on the Pi via a network connection. I used NOOBS to install some stripped-down operating systems—in this case the OSMC and LibreElec distributions—that are designed to support a single multimedia player application.

JustBoom provides software guides for several popular players and operating systems, with more on the way. On my first attempt, I tried configuring the DAC With Amp setup to work with the popular OSMC package, but I couldn't get any of the hardware settings to stick, and couldn't get any sound out of either the DAC (via the 3.5-mm socket) or the amp speaker connectors using a small passive speaker. After some frustrated poking around the JustBoom website for some missing step and general googling, I tried again using LibreElec. This OS/media player combo proved to have the advantage of a much simpler configuration process and the welcome ability to actually deliver sound out of all the various connectors.

Sticking with LibreElec, I tried the Amp HAT with some much larger passive speakers provided by our audiophile executive editor. Sound from a sample set of high-resolution music files also provided by him (including classical and rock music) was soon pouring out over an impressive range of volumes. At the very loudest point of one rock song, the Pi spontaneously rebooted, most likely because of the anemic 800-milliampere power supply I was using. Nonetheless, even with this supply, it was possible to play music at a level well beyond the comfort threshold for an indoor setup.

Summoning *Spectrum*'s executive editor for his final verdict, I cycled through the various configurations, including using active speakers to compare the Pi's native analog output to the DAC's. Zorpette was pleasantly surprised by the quality of the output, which he feels can stand its ground with even much pricier home DACs or amplifiers.

—STEPHEN CASS

RESOURCES_EDUCATION

UCONN'S OUTREACH TO K–12 GIRLS PAYS OFF

A CASE STUDY IN DOUBLING FEMALE ENGINEERING ENROLLMENT

CHRISTOPHER LAROSA/UNIVERSITY OF CONNECTICUT

LEADING BY EXAMPLE: The University of Connecticut connects girls over a range of ages with female engineering undergraduates.

The lack of women in science, technology, engineering, and mathematics (STEM) programs continues to be a concern of organizations such as the IEEE and the U.S. National Academy of Engineering. Addressing this problem means encouraging more girls to study the field at degree level: More than four out of five engineering majors in most colleges and universities in the United States are men. While a handful of programs (like at Yale, MIT, and the Franklin W. Olin College of Engineering) are already at or near 50/50 parity, the vast majority of engineering programs have a long way to go.

However, the University of Connecticut's engineering program has seen its female enrollment more than double since 2009, from 322 students in 2009 to 688 in 2015. (This translates to 28 percent of the engineering majors in the class of 2020.) The school attributes its rising female enrollment numbers to a suite of programs aimed at girls across a range of ages.

"We have a whole variety of outreach programs that span the whole K–12 space," says Daniel Burkey, associate dean for undergraduate education and diversity in the university's school of engineering. "Programs that really try and go out and speak to underrepresented populations at all levels precollegiately in order to show them what engineers do, show them positive role models in engineering that look like they do, show them how engineering makes a positive impact on the world socially and environmentally."

UConn started a local chapter of a program called Engineering Ambassadors in 2010. Through EA, student volunteers from the engineering program—many of them women—go out to primary and middle schools around Connecticut.

For eighth-grade girls, UConn engineering hosts a twice-a-year conference called Multiply Your Options. MYO, Burkey says, offers a dozen hands-on workshops led by women engineering undergraduates and a "guess my occupation" game featuring recent female engineering grads. An MYO event last October hosted an estimated 220 eighth-grade girls, Burkey says, adding that much of the impact comes from "peer mentoring, being able to talk to people not that much older than them and seeing that it's not scary and they're capable of doing that."

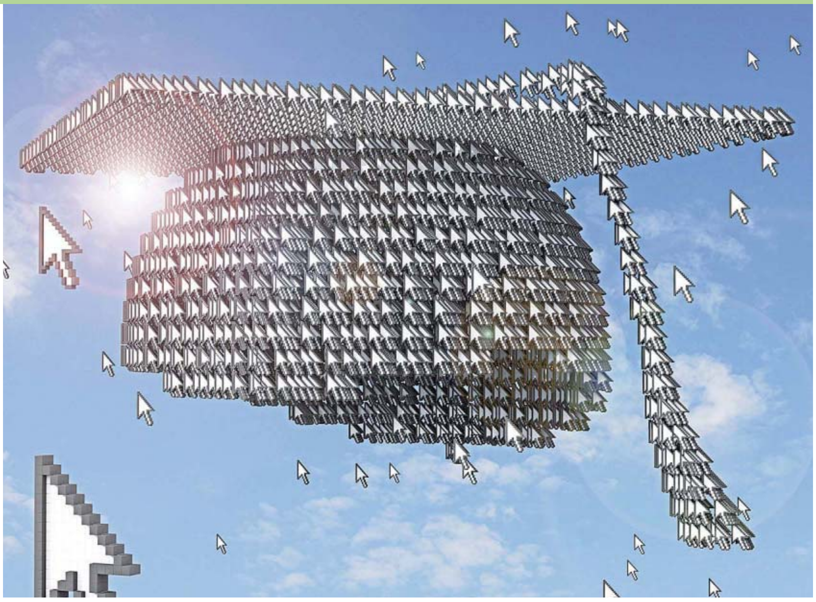
For high school students, UConn engineering last year started a program called Women in Engineering Day. This program allows participants to spend the day with current and former engineering students and faculty. In 2015, Burkey says, 120 high school girls participated, including 89 seniors. And of those seniors, 30 enrolled in UConn's school of engineering.

Shasha Graves, a junior in biomedical engineering at UConn, is active in multiple engineering outreach programs at the school as well as its section of the Society of Women Engineers. "One of the best ways to encourage and assist women in engineering anywhere is by creating a supportive environment, one where students are encouraged to be themselves and help one another succeed rather than compete with one another," she says. "Engineering is about solving problems, and the more diverse the field is the better shot we'll have at solving these problems, because everyone brings a unique perspective," says Graves.

Phoebe Szarek, now the president of UConn's Engineering Ambassadors chapter, certainly took away a similar lesson when she was an eighth grader participating in UConn's MYO program. "This was the first time that I had attended anything specifically aimed towards girls interested in STEM," Szarek says. "And this made me begin to think that maybe girls in STEM was something important, different, and special." —**MARK ANDERSON**

RESOURCES_EDUCATION

MOOCS COME BACK TO EARTH
PEDAGOGIC REALITY IS REPLACING EARLY HYPE



GETTY IMAGES

In 2011, Stanford computer scientists Sebastian Thrun and Peter Norvig came up with the bright idea of streaming their robotics lectures on demand over the Internet, and letting anyone sign up and participate in the coursework. When more than 160,000 enrolled, the professors thought they had a tiger by the tail. The MOOC—massive open online course—had arrived. To date, about 58 million people have signed up for MOOCs.

Thrun promptly cofounded Udacity to commercialize MOOCs. He predicted that in 50 years, streaming lectures would so subvert face-to-face education that only 10 higher-education institutions would remain. Campuses would become obsolete, replaced by "star" faculty streaming to screens all over the world.

MOOC pioneers were therefore stunned when their courses didn't perform anything like they had expected. At first, the average completion rate for MOOCs was less than 7 percent. Completion rates have since gone up a bit, to a median of about 12.6 percent, although there's considerable variation. While a number of factors contribute to the completion rate, most of those who did finish a MOOC were accomplished learners, many with advanced degrees My own observation, at New York University, is that students who have to pay a fee to enroll tend to be more committed to finishing.

What accounts for MOOCs' modest performance? While the technological solution they devised was novel, most MOOC innovators were unfamiliar with key educational trends. The first MOOCs replicated the standard lecture, an uninspiring teaching style, but one with which the computer scientists were familiar. Most MOOC founders were unaware of the pedagogical revolution already under way: The traditional university lecture was being rejected by many scholars, practitioners, and, most tellingly, tech-savvy students. MOOC advocates also failed to appreciate the existing body of knowledge about learning online, built over the last couple of decades by adventurous faculty who were attracted to online teaching for its innovative potential, such as peer-to-peer learning, virtual teamwork, and interactive exercises. These modes of instruction, known collectively as "active" learning, encourage student engagement rather than passive listening in lectures.

The impact of active learning can be significant. In a 2014 meta-analysis published in *Proceedings of the National Academy of Sciences*, researchers looked at 225 studies in which standard lectures were compared with active learning for undergraduate science, math, and engineering. Average test scores went up about 6 percent in active-learning sections, while students in traditional lecture classes were 1.5 times as likely to fail as their peers in active-learning classes.

Even lectures by star faculty were no match for active-learning sections taught by novice instructors: Students still performed better in active classes. "We've yet to see any evidence that celebrated lecturers can help students more than even first-generation active learning does," Scott Freeman, the lead author of the study, told *Wired*.

Unfortunately, early MOOCs failed to incorporate active learning approaches or any of the other innovations in teaching and learning common in other online courses. The three principal MOOC providers—Coursera, edX, and Udacity—wandered into a territory they thought was uninhabited. It's telling that in their latest offerings, these vendors have introduced a number of active-learning innovations.

To be sure, MOOCs have been wildly successful in giving millions of people all over the world access to a wide range of subjects presented by eminent scholars at the world's elite schools. Some courses attract so many students that a 7 percent completion rate still translates into thousands of students finishing.

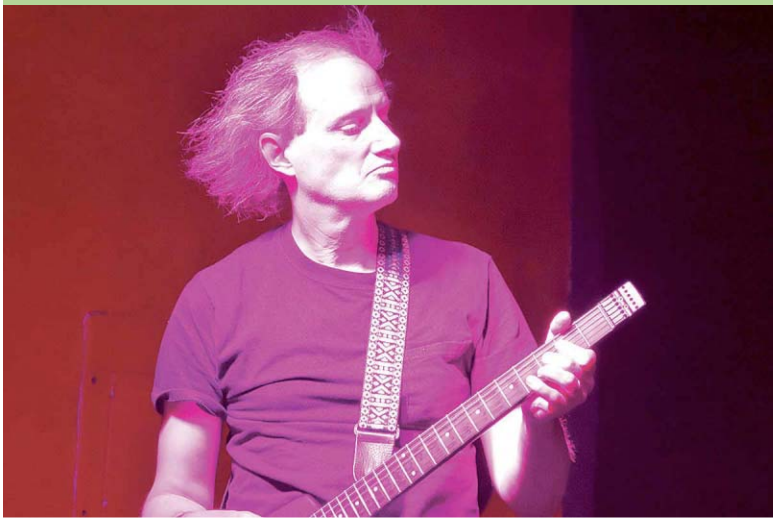
But MOOC pioneers were presumptuous to imagine they could topple the university. They erroneously assumed they could open the minds of millions who were unprepared to tackle sophisticated curricula. MOOCs will never sweep away face-to-face classrooms, nor can they take the place of more intensive and intimate online degree programs. The real contribution of MOOCs is likely to be much more modest, as yet another digital education option. **—ROBERT UBELL**

An extended version of this article is online.

RESOURCES_GEEK LIFE

DIGITAL MUSIC'S LANGUAGE MAESTRO

MILLER PUCKETTE CREATED PD AND MAX, THE CODE BEHIND MANY COMPOSITIONS



THE META COMPOSER: As well as composing and performing his own electronic music, Miller Puckette gave software tools to many others.

MILES ROSEMIRE

Experimental musicians and sound artists frequently incorporate otherworldly drones, warbles, and never-before-heard noises into their work. Yet as different as these compositions are, many of them have something in common: a pair of programming languages written by University of California, San Diego, computer musician Miller Puckette.

Those languages are Max and Pure Data, or Pd. They serve as popular all-purpose tools for composing pieces or building musical instruments, either wholly synthetic or physical-digital hybrids. These languages, now several decades old, have nourished active communities of avant-garde musicians, who rely on the software's unique visual interfaces for their performances and compositions.

Puckette wrote Max in 1988 while working at the Institute for Research and Coordination in Acoustics/Music, or IRCAM, in Paris. Creating pieces of music with a computer back then was tedious—each work required creating an all-new program. "What I was trying to do was make software that, with modifications, would be mostly reusable from one piece of music to the next," he says. The language was an instant hit.

Max programs are called patches and look like sprawling flowcharts. Composers drag boxes, called objects, around the screen and link them together with connectors, called patch cords. (The term *patch cords* came by way of the earlier age of analog modular synthesizers where oscillators, filters, and other modules were physically interconnected by such cables.)

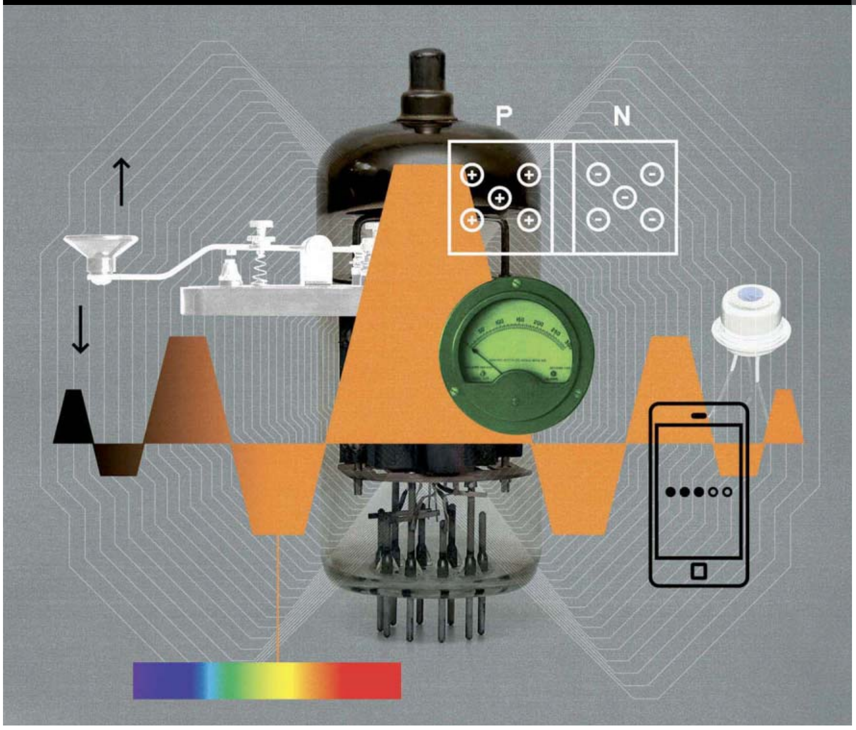
Inputs to objects are called messages and they can be sounds, keyboard presses, or any other piece of digital information. Objects process messages according to rules dictated by the programmer/composer and send outgoing messages to other objects or output devices like loudspeakers or lights. Programmers can export patches as stand-alone pieces of software or integrate them with other audio programs, synthesizers, and controllers.

Max was not created under an open source license, and Puckette discovered he could not always make the modifications he wanted. "As a minion, all the code I wrote belonged to IRCAM," he says. So sometime around 1996, he started over with Pd. (No one knows when Pd really started, Puckette says, but the Pd community considers 2016 the software's 20th year.)

Pd is similar to Max and works on the same basic principle of combining objects by passing messages. "I changed a lot of things I thought were bad about Max," says Puckette. He made Pd faster, lighter, and more stable.

Puckette stays active in the Pd community. He "has started doing a lot more performance himself," says Kerry Hagan, a researcher at the University of Limerick's Digital Media & Arts Research Centre, in Ireland. She was Puckette's graduate student around the time he first drafted Pd, and today she teaches both Pd and Max to her students. Puckette and Hagan often pass patches back and forth, experimenting with new ways to produce sounds using the laws of mathematics and probability. Hagan incorporates their combined efforts into her own performances, she says: "Miller's programming and math, and my art."

Pd users now gather annually at PdCon to showcase their patches, offer suggestions for improvement, and perform. Puckette and another former student, New York University computer music assistant professor Jaime E. Oliver La Rosa, put on a show together to close out the 2016 gathering last November. La Rosa sat over his novel instrument, called the Mano Controller, which looks like a large felt-covered music stand, next to a large iMac running his Pd patch. A camera above the stand converts the player's hand positions into patch inputs. Puckette held a piece of wired plastic in one hand and waved the other above a box-shaped controller on a chair. He contorted his fingers into several configurations, and the shapes acted as the input to another Pd patch. A cacophony of processed drones, clicks, and clanks that sounded like a jungle inhabited solely by robots spewed forth for several minutes. Both men giggled at the performance's conclusion. **—RYAN F. MANDELBAUM**



THANK MAXWELL FOR CELLPHONES

NOW THAT THE WORLD HAS BECOME ADDICTED to portable electronics, billions of people have come to see the companies providing these gadgets as the most innovative, and the people who head those companies as the most exalted, of all time. "Genius" is a starter category in this discussion. ● But clever and appealing though today's electronic gadgets may be, to the historian they are nothing but the inevitable fifth-order elaborations of two fundamental ideas: electromagnetic radiation, the theory of which was formulated by James Clerk Maxwell in the 1860s, and miniaturized fabrication, which followed Richard Feynman's 1959 dictum that "there's plenty of room at the bottom." ● Maxwell was a true genius. The history of science offers few examples of work as brilliant as unifying electricity, magnetism, and light as aspects of a single phenomenon: electromagnetic waves. As Max Planck put it, "in doing so he achieved greatness unequalled." ● In late 1879 and early 1880, David Edward Hughes actually transmitted and received those invisible signals but did not publish his results. And in 1883, Thomas Edison also came very close to the actual use of such waves with his patent for an apparatus "showing conductivity of continuous currents through high vacuo." He displayed it in 1884 at the International Electrical Exhibition in Philadelphia, then abandoned it as a mere curiosity. ● That is why the second-order elaboration of electromagnetism came only between 1886 and 1888, when Heinrich Hertz deliberately generated and received electromagnetic waves whose frequencies he accurately placed "in a position intermediate between the acoustic oscillations of ponderable bodies and the light-oscillations of the ether."

The third-order elaboration began with the first broadcasts, by Oliver J. Lodge and Alexander S. Popov, in 1894 and 1895, and it continued with the first transatlantic transmission, by Guglielmo Marconi in 1901; the first long-distance transmission of voice and music, by Reginald A. Fessenden in 1906; and the invention of vacuum tubes–John A. Fleming's diode in 1904, Greenleaf W. Pickard's point-contact diode (cat's whisker) in 1906, and Lee de Forest's triode in 1907.

By 1910, we thus had in place all the fundamental technical requirements for commercial radio, and even perhaps for primitive television. And in the 1920s we came close to solid-state electronics: Julius Edgar Lilienfeld filed to patent the concept of a field-effect transistor in 1925 in Canada and a year later in the United States. In 1934, Oskar Heil patented a similar device, and in 1939 Russell Ohl discovered the *p-n* junction. Consequently, even the Bell Labs Memorial site acknowledged that its scientists' much-publicized invention of the transistor in 1947 was in fact a reinvention.

Solid-state devices opened the gates of the fourth-order elaboration, beginning with integrated circuits–a patent for which was filed in 1959–followed by microprocessors in 1971 and 1972. Bell Laboratories, RCA, Fairchild Semiconductor, Texas Instruments, and Intel were the leading enablers of this wave.

The fifth-order elaboration–the one that now receives all the adulation of the texting and tweeting masses–followed inexorably, as microprocessors proliferated in mobile communications. Once again, this elaboration was built on basic advances made long before.

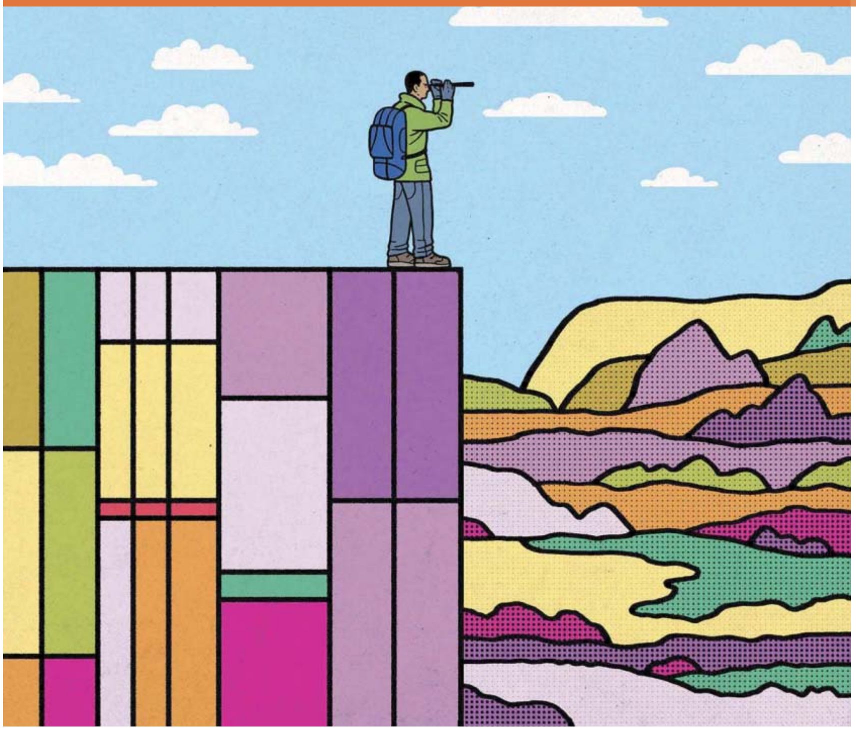
As I pass the zombielike figures on the street, oblivious to anything but their cellphone screens, I wonder how many of them know that the most fundamental advances enabling their addictions came not from Nokia, Apple, Google, Samsung, or LG. These companies' innovations are certainly admirable, but they amount only to adding a few fancy upper floors to a magnificent edifice whose foundations were laid by Maxwell 152 years ago and whose structure depends on decades-old advances that made it possible to build electronics devices ever smaller.

Not that that's a small matter. ■

ILLUSTRATION BY Stuart Bradford

REFLECTIONS_BY ROBERT W. LUCKY

OPINION



THE MILLIMETER WIRELESS FRONTIER

THERE IS AN ETERNAL QUALITY TO HOW TECHNOLOGY EVOLVES. As existing areas get overworked, new frontiers open up at the fringes. Then innovators rush in to occupy the new territory before it, in turn, becomes overworked. There is an example of such a frontier today in wireless communications. IEEE's 5G wireless initiative has the goal of serving many more users with much higher transmission speeds. But with the existing cellular bands tightly packed, where does all the required additional network capacity come from? ● In contrast with the traditional radio-spectrum management view of scarce capacity, where a finite amount of spectrum must be divided up among users, communication theorists see wireless capacity as virtually unlimited. Capacity can be increased indefinitely by going to ever smaller cells and higher frequencies that offer more bandwidth, while greater efficiency can be achieved with advanced signal processing and new spectrum-sharing policies. Among all these approaches, the greatest immediate impact would be achieved by moving to the higher frequencies in the millimeter range—the region of 30 to 300 gigahertz, where bandwidth is available and plentiful. ● But in many ways, millimeter-wave wireless truly is a frontier. Today the millimeter band is largely uninhabited and inhospitable, as signals using these wavelengths run up against difficult propagation problems. Even when signals travel through free space, attenuation increases with frequency, so usable path lengths for millimeter waves are short, roughly 100 to 200 meters. Such distances could be accommodated with the smaller cell sizes envisioned in 5G, but there are numerous other impediments. Buildings and the objects in and around them, including people, block the signal. Rain and foliage further attenuate millimeter waves, and diffraction—which can bend longer wavelengths around occluding objects—is far less effective. Even surfaces that might be conveniently nicely reflective at longer wavelengths appear rougher to millimeter waves, and so diffuse the signal.

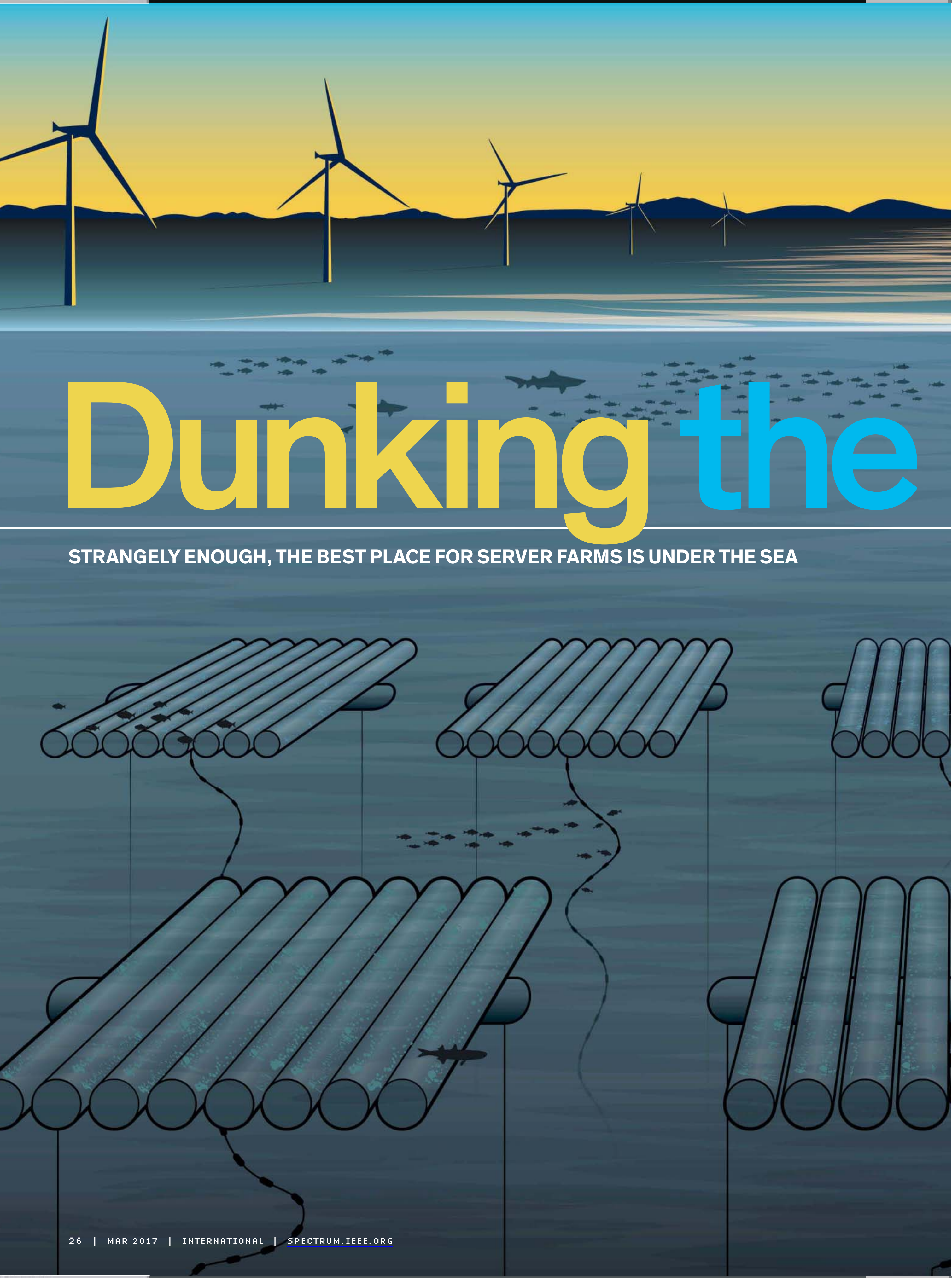
So there may be gold in that frontier, but it is going to be very difficult to mine. Nevertheless, you never know until you try. I'm reminded of Marconi's successful transatlantic transmission in 1901, when physicists insisted that the signal would fly off into space. Recently a team at NYU has been experimenting with millimeter-wave transmission within the urban canyons of New York City. Like the physicists of yesteryear, I would have said that this would never work. But the data show otherwise. They demonstrated a surprising amount of coverage despite the buildings, pedestrian and vehicular traffic, and general chaos typical of dense cities. Granted, there are a number of holes in the coverage, but initial results are encouraging.

When I expressed some surprise at these findings, someone pointed out to me that in its line-of-sight dependence, millimeter-wave propagation might be likened to that of visible light, and that the nighttime world isn't as dark as might be expected. Taking this analogy to heart, I prowled my house on a dark night, the lone source of illumination a weak light at the end of a long corridor. I discovered dim light in unexpected places. I wondered: How did the light get here? Even so, nearby rooms might be caves of darkness. All the while, I was conscious of the strong Wi-Fi signal that followed me everywhere I went.

So the millimeter-wave frontier is going to be a difficult one, but we engineers are good at this kind of challenge, and we're not without tools. For one thing, at these small wavelengths, we can build postage-stamp-size phased-array antennas, and high-speed electronics allow us to use advanced techniques that have been pioneered at longer wavelengths.

All this sophisticated technology so we will be able to view 3D video of cats while walking down a busy city street. Or, hopefully, some other use. ■

ILLUSTRATION BY Dan Page



Dunking the

STRANGELY ENOUGH, THE BEST PLACE FOR SERVER FARMS IS UNDER THE SEA

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Data Center

By **Ben Cutler**, **Spencer Fowers**, **Jeffrey Kramer** & **Eric Peterson**

Illustration by **MCKIBILLO**

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

WHEN SEAN JAMES, who works on data-center technology for Microsoft, suggested that the company put server farms entirely underwater, his colleagues were a bit dubious. But for James, who had earlier served on board a submarine for the U.S. Navy, submerging whole data centers beneath the waves made perfect sense. ● This tactic, he argued, would not only limit the cost of cooling the machines—an enormous expense for many data-center operators—but it could also reduce construction costs, make it easier to power these facilities with renewable energy, and even improve their performance.

Together with Todd Rawlings, another Microsoft engineer, James circulated an internal white paper promoting the concept. It explained how building data centers underwater could help Microsoft and other cloud providers manage today's phenomenal growth in an environmentally sustainable way.

At many large companies, such outlandish ideas might have died a quiet death. But Microsoft researchers have a history of tackling challenges of vital importance to the company in innovative ways, even if the required work is far outside of Microsoft's core expertise. The key is to assemble engineering teams by uniting Microsoft employees with colleagues from partner companies.

The four of us formed the core of just such a team, one charged with testing James's far-out idea. So in August of 2014, we started to organize what soon came to be called Project Natick, dubbed that for no particular reason other than that our research group likes to name projects after cities in Massachusetts. And just 12 months later, we had a prototype serving up data from beneath the Pacific Ocean.

Project Natick had no shortage of hurdles to overcome. The first, of course, was keeping the inside of its big steel container dry. Another was figuring out the best way to use the surrounding seawater to cool the servers inside. And finally there was the matter of how to deal with the barnacles and other forms of sea life that would inevitably cover a submerged vessel—a phenomenon that should be familiar to anyone who has ever kept a boat in the water for an extended period. Clingy crustaceans and such would be a challenge because they could interfere with the transfer of heat from the servers to the surrounding water. These issues daunted us at first, but we solved them one by one, often drawing on time-tested solutions from the marine industry.

But why go to all this trouble? Sure, cooling computers with seawater would lower the air-conditioning bill and could improve operations in other ways, too, but submerging a data center comes with some obvious costs and inconveniences. Does trying to put thousands of computers beneath the sea really make sense? We think it does, for several reasons.

ALL PHOTOGRAPHS: MICROSOFT

IN THE PINK: Equipment submerged in the ocean for long periods is prone to encrustation with marine organisms. Despite such biofouling, the authors' prototype data-center pod operated flawlessly.

For one, it would offer a company like ours the ability to quickly target capacity where and when it is needed. Corporate planners would be freed from the burden of having to build these facilities long before they are actually required in anticipation of later demand. For an industry that spends billions of dollars a year constructing ever-increasing numbers of data centers, quick response time could provide enormous cost savings.

The reason underwater data centers could be built more quickly than land-based ones is easy enough to understand. Today, the construction of each such installation is unique. The equipment might be the same, but building codes, taxes, climate, workforce, electricity supply, and network connectivity are different everywhere. And those variables affect how long construction takes. We also observe their effects in the performance of our facilities, where otherwise identical equipment exhibits different levels of reliability depending on where it is located.

As we see it, a Natick site would be made up of a collection of "pods"–steel cylinders that would each contain possibly several thousand servers. Together they'd make up an underwater data center, which would be located within a few kilometers of the coast and placed between 50 and 200 meters below the surface. The pods could either float above the seabed at some intermediate depth, moored by cables to the ocean floor, or they could rest on the seabed itself.

Once we deploy a data-center pod, it would stay in place until it's time to retire the set of servers it contains. Or perhaps market conditions would change, and we'd decide to move it somewhere else. This is a true "lights out" environment, meaning that the system's managers would work remotely, with no one to fix things or change out parts for the operational life of the pod.

Now imagine applying just-in-time manufacturing to this concept. The pods could be constructed in a factory, provisioned with servers, and made ready to ship anywhere in the world. Unlike the case on land, the ocean provides a very uniform environment wherever you are. So no customization of the pods would be needed, and we could install them quickly anywhere that computing capacity was in short supply, incrementally increasing the size of an underwater installation to meet capacity requirements as they grew. Our goal for Natick is to be able to get data centers up and running, at coastal sites anywhere in the world, within 90 days from the decision to deploy.

Most new data centers are built in locations where electricity is inexpensive, the climate is reasonably cool, the land is cheap, and the facility doesn't impose on the people living nearby. The problem with this approach is that it often puts data centers far from population centers, which limits how fast the servers can respond to requests.

For interactive experiences online, these delays can be problematic. We want Web pages to load quickly and video games such as *Minecraft* or *Halo* to be snappy and lag free. In years to come, there will be more and more interaction-rich applications, including those enabled by Microsoft's HoloLens and other mixed reality/virtual reality technologies. So what you really want is for the servers to be close to the people they serve, something that rarely happens today.

It's perhaps a surprising fact that almost half the world's population lives within 100 kilometers of the sea. So placing data centers just offshore near coastal cities would put them much closer to customers than is the norm today.

If that isn't reason enough, consider the savings in cooling costs.

THREE-MONTH DIP: A data-center pod containing server racks and using heat exchangers for cooling was tested off the coast of California, where it operated for 105 days at a depth of 11 meters.

Historically, such facilities have used mechanical cooling–think home air-conditioning on steroids. This equipment typically keeps temperatures between 18 and 27 °C, but the amount of electricity consumed for cooling is sometimes almost as much as that used by the computers themselves.

More recently, many data-center operators have moved to free-air cooling, which means that rather than chilling the air mechanically, they simply use outside air. This is far cheaper, with a cooling overhead of just 10 to 30 percent, but it means the computers are subject to outside air temperatures, which can get quite warm in some locations. It also often means putting the centers at high latitudes, far from population centers.

What's more, these facilities can consume a lot of water. That's because they often use evaporation to cool the air somewhat before blowing it over the servers. This can be a problem in areas subject to droughts, such as California, or where a growing population depletes the local aquifers, as is happening in many developing countries. Even if water is abundant, adding it in the air makes the electronic equipment more prone to corrosion.

Our Natick architecture sidesteps all these problems. The interior of the data-center pod consists of standard computer racks with attached heat exchangers, which transfer the heat from the air to some liquid, likely ordinary water. That liquid is then pumped to heat exchangers on the outside of the pod, which in turn transfer the heat to the surrounding ocean. The cooled transfer liquid then returns to the internal heat exchangers to repeat the cycle.

Of course, the colder the surrounding ocean, the better this scheme will work. To get access to chilly seawater even during the summer or in the tropics, you need only put the pods sufficiently deep. For example, at 200 meters' depth off the east coast of Florida, the water remains below 15 °C all year round.

Our tests with a prototype Natick pod, dubbed the "Leona Philpot" (named for an Xbox game character), began in August 2015. We submerged it at just 11 meters' depth in the Pacific near San Luis Obispo, Calif., where the water ranged between 14 and 18 °C.

Over the course of this 105-day experiment, we showed that we could keep the submerged computers at temperatures that were at least as cold as mechanical cooling can achieve and with even lower energy overhead than the free-air approach—just 3 percent. That energy-overhead value is lower than any production or experimental data center of which we are aware.

Because there was no need to provide an on-site staff with lights to see, air to breathe, parking spaces to fight over, or big red buttons to press in case of emergency, we made the atmosphere in the data-center pod oxygen free. (Our employees managed the prototype Natick pod from the comfort of their Microsoft offices.) We also removed all water vapor and dust. That made for a very benign environment for the electronics, minimizing problems with heat dissipation and connector corrosion.

Microsoft is committed to protecting the environment. In satisfying its electricity needs, for example, the company uses renewable sources as much as possible. To the extent that it can't do that, it purchases carbon offsets. Consistent with that philosophy, we are looking to deploy our future underwater data centers near offshore sources of renewable energy—be it an offshore wind farm or some marine-based form of power generation that exploits the force of tides, waves, or currents.

These sources of energy are typically plentiful offshore, which means we should be able to match where people are with where we can place our energy-efficient underwater equipment and where we would have access to lots of green energy. Much as data centers today sometimes act as anchor tenants for new land-based renewable-energy farms, the same may hold true for marine energy farms in the future.

Another factor to consider is that conventionally generated electricity is not always easily available, particularly in the developing world. For example, 70 percent of the population of sub-Saharan Africa has no access to an electric grid. So if you want to build a data center to bring cloud services closer to such a population, you'd probably need to provide electricity for it, too.

Typically, electricity is carried long distances at 100,000 volts or higher, but ultimately servers use the same kinds of low voltages as your PC does. To drop the grid power to a voltage that the servers can consume generally requires three separate pieces of equipment. You also need backup generators and banks of batteries in case grid power fails.

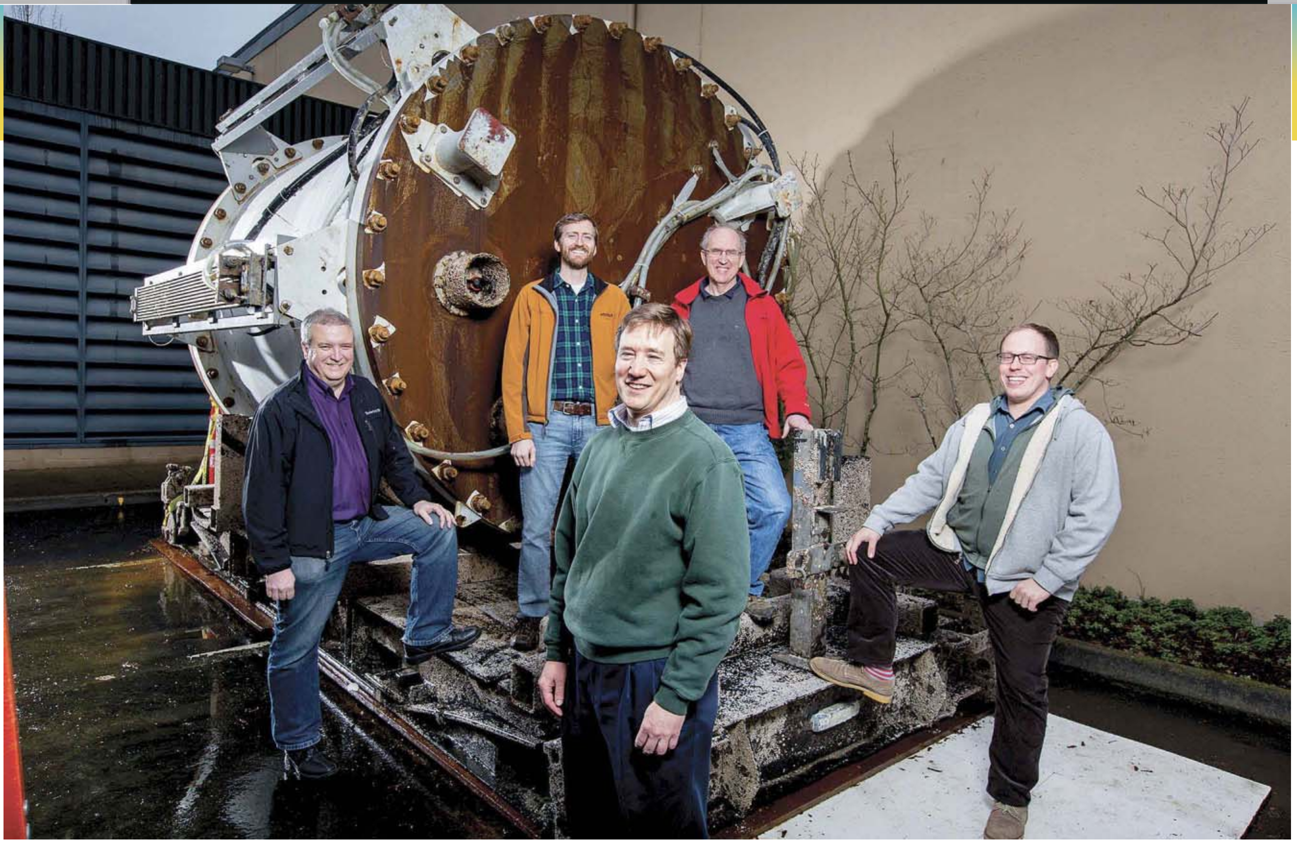
Locating underwater data centers alongside offshore sources of power would allow engineers to simplify things. First, by generating power at voltages closer to what the servers require, we could eliminate some of the voltage conversions. Second, by powering the computers with a collection of independent wind or marine turbines, we could automatically build in redundancy. This would reduce both electrical losses and the capital cost (and complexity) associated with the usual data-center architecture, which is designed to protect against failure of the local power grid.

An added benefit of this approach is that the only real impact on the land is a fiber-optic cable or two for carrying data.

The first question everyone asks when we tell them about this idea is: How will you keep the electronics dry? The truth is that keeping things dry isn't hard. The marine industry has been keeping equipment dry in the ocean since long before computers even existed, often in far more challenging contexts than anything we have done or plan to do.

The second question—one we asked ourselves early on—is how to cool the computers most efficiently. We explored a range of exotic approaches, including the use of special dielectric liquids and phase-change materials as well as unusual heat-transfer media such as high-pressure helium gas and supercritical carbon dioxide. While such approaches have their benefits, they raise thorny problems as well.

While we continue to investigate the use of exotic materials for cool-



GOT IT BACK: Authors Cutler [center], Peterson [left], Fowers [middle left], and Kramer [right], along with Microsoft Research's director of special projects, Norm Whittaker [middle right], celebrate a successful test.

ing, for the near term we see no real need. Natick's freshwater plumbing and radiator-like heat exchangers provide a very economical and efficient cooling mechanism, one that works just fine with standard servers.

A more pertinent issue, as we see it, is that an underwater data center will attract sea life, in effect forming an artificial reef. This process of colonization by marine organisms, called biofouling, starts with single-celled creatures, which are followed by somewhat larger organisms that feed on those cells, and so on up the food chain.

When we deployed our Natick prototype, crabs and fish began to gather around the vessel within 24 hours. We were delighted to have created a home for those creatures, so one of our main design considerations was how to maintain that habitat while not impeding the pod's ability to keep its computers cool.

In particular, we knew that biofouling on external heat exchangers would disrupt the flow of heat from those surfaces. So we explored the use of various antifouling materials and coatings—even active deterrents involving sound or ultraviolet light—in hopes of making it difficult for life to take hold. Although it's possible to physically clean the heat exchangers, relying on such interventions would be unwise, given our goal to keep operations as simple as possible.

Thankfully, the heat exchangers on our Natick pod remained clean during its first deployment, despite it being in a very challenging setting (shallow and close to shore, where ocean life is most abundant). But biofouling remains an area of active research, one that we continue to study with a focus on approaches that won't harm the marine environment.

The biggest concern we had by far during our test deployment was that equipment would break. After all, we couldn't send a tech to some server rack to swap out a bad hard drive or network card. Responses to hardware failures had to be made remotely or autonomously. Even in Microsoft's data centers today, we and others have been working to increase our ability to detect and address failures without human intervention. Those same techniques and expertise will be applied to Natick pods of the future.

How about security? Is your data safe from cyber or physical theft if it's underwater? Absolutely. A Natick site would provide the same encryption and other security guarantees of a land-based Microsoft data center. While no people would be physically present, sensors would give a Natick pod an excellent awareness of its surroundings, including the presence of any unexpected visitors.

You might wonder whether the heat from a submerged data center would be harmful to the local marine environment. Not likely. Any heat generated by a Natick pod would rapidly be mixed with cool water and carried away by the currents. The water just meters downstream of a Natick vessel would get a few thousandths of a degree warmer at most.

So the environmental impact would be very modest. That's important, because the future is bound to see a lot more data centers get built. If we have our way, though, people won't actually see many of them, because they'll be doing their jobs deep underwater. ■

↗ POST YOUR COMMENTS at http://spectrum.ieee.org/underseadata0317

Deep Learning Reinvents the Hearing Aid

Finally, wearers can pick out a voice in a crowded room

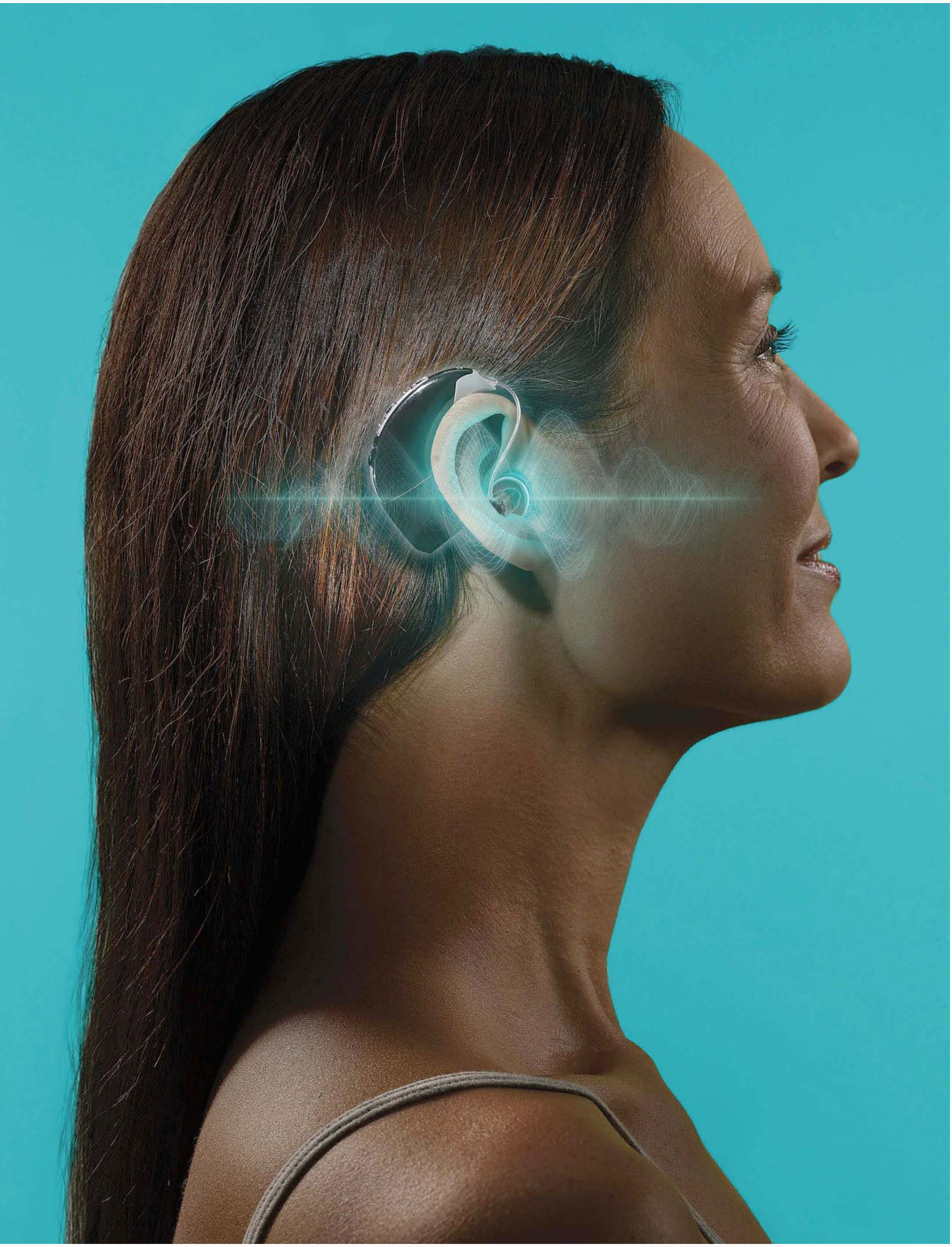
By DeLiang Wang

MY MOTHER BEGAN TO LOSE HER HEARING while I was away at college. I would return home to share what I'd learned, and she would lean in to hear. Soon it became difficult for her to hold a conversation if more than one person spoke at a time. Now, even with a hearing aid, she struggles to distinguish the sounds of each voice. When my family visits for dinner, she still pleads with us to speak in turn.

My mother's hardship reflects a classic problem for hearing aid manufacturers. The human auditory system can naturally pick out a voice in a crowded room, but creating a hearing aid that mimics that ability has stumped signal processing specialists, artificial intelligence experts, and audiologists for decades. British cognitive scientist Colin Cherry first dubbed this the "cocktail party problem" in 1953.

More than six decades later, less than 25 percent of people who need a hearing aid actually use one. The greatest frustration among potential users is that a hearing aid cannot distinguish between, for example, a voice and the sound of a passing car if those sounds occur at the same time. The device cranks up the volume on both, creating an incoherent din.

It's time we solve this problem. To produce a better experience for hearing aid wearers, my lab at Ohio State University, in Columbus, recently applied machine learning based on deep neural networks to the task of segregating sounds. We have tested multiple versions of a digital filter that not only amplifies sound but can also isolate speech



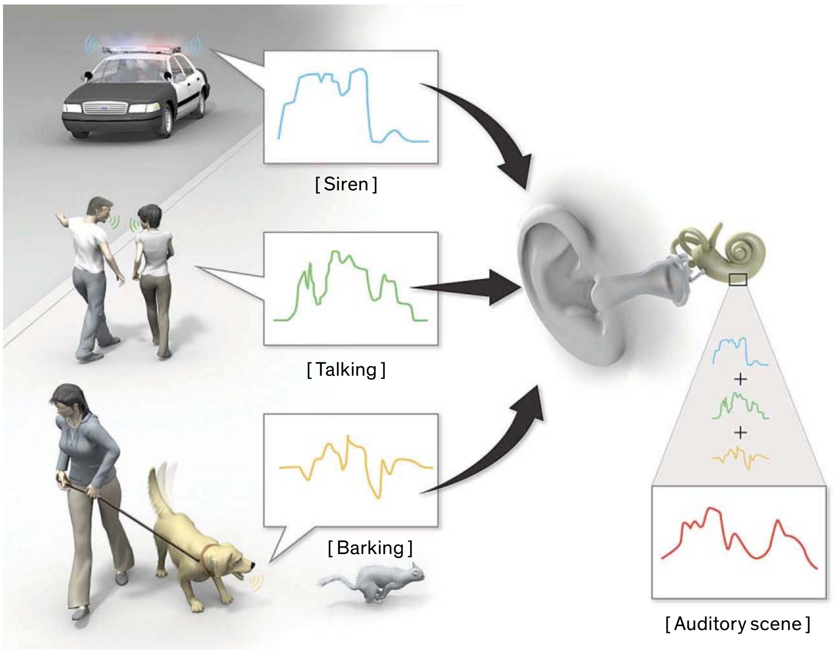
IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

from background noise and automatically adjust the volumes of each separately.

We believe this approach can ultimately restore a hearing-impaired person's comprehension to match–or even exceed–that of someone with normal hearing. In fact, one of our early models boosted, from 10 to 90 percent, the ability of some subjects to understand spoken words obscured by noise. Because it's not necessary for listeners to understand every word in a phrase to gather its meaning, this improvement frequently meant the difference between comprehending a sentence or not.

Without a better hearing aid, the world's hearing will get worse. The World Health Organization estimates that 15 percent of adults, or roughly 766 million people, suffer from hearing loss. That number is rising as the population expands and the proportion of older adults becomes larger. And the potential market for an advanced hearing aid isn't limited to people with hearing loss. Developers could use the technique to improve smartphone speech recognition. Employers could use it to help workers on noisy factory floors, and militaries could equip soldiers to hear one another through the noisy chaos of warfare.

Satisfying all those new customers, though, means finding a way to put the cocktail party problem behind us. At last, deep neural networks are pointing the way forward.



A NOISY WORLD: Thanks in part to its odd shape, the human ear captures many sound streams at once. A stream is all the sound waves that emanate from a single source, such as a dog. Together, these streams make up an auditory scene (barking + siren + talking).

For decades, electrical and computer engineers tried and failed to achieve speech isolation through signal processing. The most popular approach has been to use a voice-activity detector to identify gaps between people's utterances as they speak. In this approach, the system designates the sounds captured within those gaps as "noise." Then, an algorithm subtracts the noise from the original recording–leaving, ideally, noise-free speech.

Unfortunately this technique, known as spectral subtraction, is notorious for removing too much speech or too little noise. Too often, what results is an unpleasant artifact (called musical noise) that makes the audio sound as if it were recorded underwater. The problems are so serious that even after many years of development, this method does little or nothing to improve people's ability to recognize speech in noisy environments.

I realized we had to take a different approach. We began with a theory from Albert Bregman, a psychologist at McGill University, in Montreal, who proposed in 1990 that the human auditory system organizes sounds into distinct streams. A stream essentially corresponds to sound emitted from a single source, such as a nearby friend. Each sound stream is unique in its pitch, volume, and the direction from which it comes.

Altogether, many streams–such as that friend speaking over the roar of a hockey game–make up what Bregman calls an "auditory scene." If sounds share the same frequency band at the same time, the loudest sound in a scene overpowers the others–a useful principle known as auditory masking. For example, someone may not notice a clock ticking in the corner of the room if rain is pattering on the roof. This principle, among others, is exploited in MP3 files to shrink the files to one-tenth of their original size by removing masked sounds (such as the ticking clock, in this case) without users noticing the omission.

Recalling Bregman's work, we wondered if we could build a filter to determine whether one sound stream dominates others at a given moment inside a specific frequency band. Psychoacousticians, who study sound perception, divide the average human's hearing range into about two dozen bands between 20 hertz and 20,000 Hz. We wanted a filter to tell us whether a sound stream containing speech or noise was stronger at certain times within these bands, as a first step toward separating the two.

My lab was the first, in 2001, to design such a filter, which labels sound streams as dominated by either speech or noise. With this filter, we would later develop a machine-learning program that separates speech from other sounds based on a few distinguishing features, such as amplitude (loudness), harmonic structure (the particular arrangement of tones), and onset (when a particular sound begins relative to others).

This original filter was what we called the ideal binary mask. It labels noise and speech that it finds within segments of sound called time-frequency units,

PREVIOUS PAGE: DAN SAELINGER/TRUNK ARCHIVE

ILLUSTRATIONS BY Emily Cooper

which designate a particular brief interval within a specific frequency band. The filter analyzes each time-frequency unit in a sample of noisy speech and marks each as either 1 or 0. It records a 1 if the "target" sound (in this case, speech) is louder than noise, and a 0 if the target sound is softer. The result is a set of 1s and 0s that represent the dominance of noise or speech within a sample. Then, the filter tosses out all units labeled 0 and reconstructs the speech from those that scored 1. To reconstruct an intelligible sentence from noisy speech, a certain percentage of time-frequency units must be labeled 1.

We began testing the ideal binary mask in 2006 with the U.S. Air Force Research Laboratory, in Ohio. Around the same time, a team from Syracuse University, in New York, independently evaluated the ideal binary mask. In those trials, the filter helped people with a hearing impairment and also listeners with normal hearing to better understand sentences mixed with noise.

We had, basically, created a speech filter that performed flawlessly in the lab. But this filter enjoyed an unrealistic advantage. By design, we had provided it with samples of speech and noise separately and then tested it using mixtures of those same samples. Because it had been given the answers (that's why it's "ideal"), the filter knew when the speech was louder than the background noise. A practical speech filter must, entirely on its own and on the fly, separate a voice from the noise in a room.

Nevertheless, the fact that the ideal binary mask dramatically improved speech comprehension for both hearing-impaired listeners and those with normal hearing had a profound implication. It demonstrated that the technique of classification, a form of supervised learning, could be employed to approximate the ideal binary mask as a way of separating speech from noise. With classification, a machine mimics human learning, in effect, by completing exercises, receiving feedback, and drawing and remembering lessons from its experiences. That's essentially the same way people learn from a young age to treat apples as a class distinct from oranges.

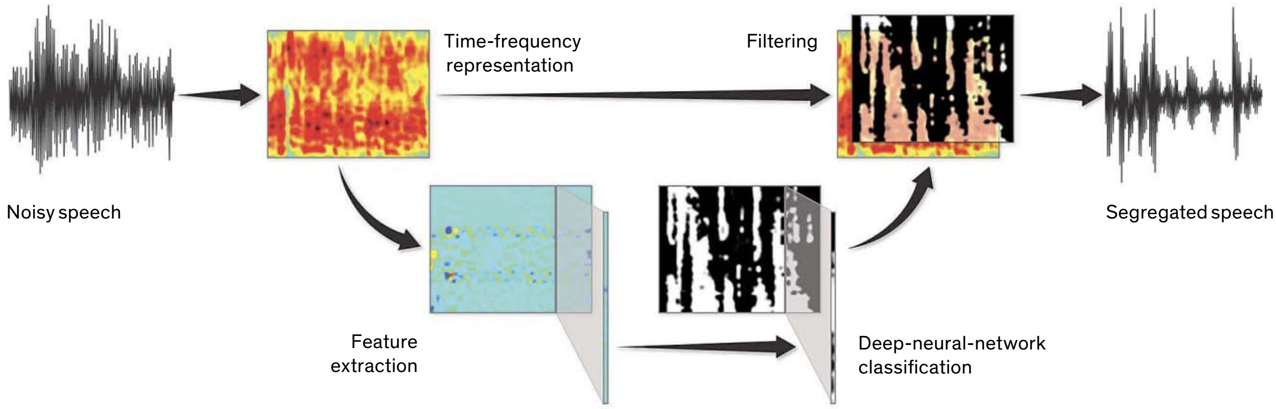
In the following years, my lab made the first attempt to approximate the ideal binary mask through classification. At about the same time we were developing our original classifier, a group at Carnegie Mellon University, in Pittsburgh, devised their own method, based on machine learning, to classify time-frequency units for another purpose: to improve automatic speech recognition. Later, a group at the University of Texas at Dallas led by the late Philipos Loizou used a different classification method. It became the first to show meaningful improvement in speech intelligibility for people with normal hearing by relying on only monaural features (as opposed to the binaural ones captured by two ears).

But these early machine-learning methods applied classification techniques that were not powerful or accurate enough to help hearing aid wearers. They could not yet handle the complex and unpredictable mixture of noises and voices that occur in the world. In order to do that, we would need something far more powerful.

Having demonstrated promising initial results with our early classification algorithms, we decided to take the next logical step—to improve the system so it could function in noisy real-world environments, and without training for specific noises and sentences. This challenge prompted us to try to do something that had never been done before: build a machine-learning program that would run on a neural network and separate speech from noise after undergoing a sophisticated training process. The program would use the ideal binary mask to guide the training of the neural network. And it worked. In a study involving 24 test subjects, we demonstrated that this program could boost the comprehension of hearing-impaired people by about 50 percent.

Basically, a neural network is a software system constructed of relatively simple elements that can achieve complex levels of processing by working together. (The system's structure is roughly modeled on how neurons and their networks work in the brain.) When presented with new examples, neural networks, like human brains, can "learn" by adjusting the weights of their connections.

CLEAN SPEECH: To separate speech from noise, a machine-learning program breaks a noisy speech sample into a collection of elements called time-frequency units. Next, it analyzes these units to extract 85 features known to distinguish speech from other sounds. Then, the program feeds the features into a deep neural network trained to classify the units as speech or not, based on past experience with similar samples. Lastly, the program applies a digital filter that tosses out all the nonspeech units to leave only separated speech.



Neural networks come in many shapes and sizes and with varying degrees of complexity. Deep neural networks are defined as having at least two "hidden" processing layers, which are not directly connected to a system's input or output. Each hidden layer refines the results fed to it by previous layers, adding in new considerations based on prior knowledge.

For example, a program designed to verify a customer's signature might begin by comparing a new signature to a sample included in a training database. However, that program also knows from its training that the new signature does not need to precisely match the original. Other layers can determine if the new signature shares certain qualities that tend to remain consistent in a person's signature, such as the angle of slant, or the failure to dot the letter *i*.

To build our own deep neural network, we began by writing algorithms to extract features that could distinguish voices from noise based on common changes in amplitude, frequency, and the modulations of each. We identified dozens of attributes that could help our program discriminate between speech and noise to some extent, and we used all 85 of them to make the algorithms as powerful as possible. Among the most important attributes we identified were the frequencies of the sounds and their intensities (loud or soft).

Next, we trained the deep neural network to use these 85 attributes to distinguish speech from noise. This training occurred in two phases: First, we set the program's parameters through unsupervised learning. This means we loaded many examples of the attributes into the program in order to prime it for the types of signals it would later have to classify on the fly.

Then we used samples of noisy speech and their corresponding results on the ideal binary mask to complete the second phase of training, which was the supervised learning. In particular, the set of 1s and 0s that make up the ideal binary mask was like an answer sheet that we used to test and improve our program's ability to separate speech from noise. For each new sample, the program would extract a set of attributes from the noisy speech. Then, after analyzing these attributes–frequencies, intensities, and so on–the filter performed a provisional classification–was it speech? was it noise?–and compared the result to what the ideal binary mask would determine in the same situation. If the result was different from the 1s and 0s within our perfect binary mask filter, we tweaked the neural network's parameters accordingly, so that the network would produce results closer to the 1s and 0s of the ideal binary mask on its next try.

To make these adjustments, we first calculated the error of the neural network, measured as the discrepancy between the ideal binary mask and the result at the neural network's final layer, which is known as the output layer. Once we computed this error, we would then use it to change the weights of the neural network's connections so that if the same classification was carried out again, the discrepancy would be reduced. The training of the neural network consisted of performing this procedure thousands of times.

One important refinement along the way was to build a second deep neural network that would be fed by the first one and fine-tune its results. While that first network had focused on labeling attributes within each individual time-frequency unit, the second network would examine the attributes of several units near a particular one. In other words, the second network provided the first network with extra context about the speech and noise it processed and further improved its classification accuracy. For example, a syllable may span many time-frequency units, but the background noise could change abruptly while it was being spoken. In our case, having contextual clues could help the program to more accurately separate speech from noise within the syllable.

At the end of the supervised training, the deep-neural-network classifier proved to be far superior to earlier methods at separating speech from noise. In fact, this algorithm was the first, of any technique relying on monaural techniques, to achieve major improvements in hearing-impaired listeners' ability to make sense of spoken phrases obscured by noise.

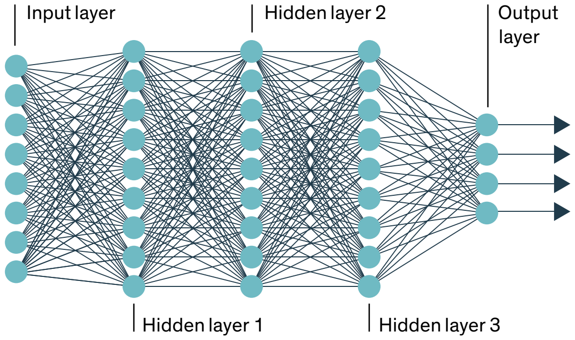
To test it with human subjects, we asked 12 hearing-impaired people and 12 with normal hearing to listen through

1880 to 1920

1921 to 1953

1984 to present

FROM TOP: HEARING AID MUSEUM; JOE HAUPT/FLICKR; SSPL/GETTY IMAGES



SMART LAYERS: A deep neural network consists of two or more processing layers in between the input layer, through which information is fed into the system [left], and the output layer, which reveals the results [right]. To improve performance, researchers can adjust the system's parameters and tweak the connections between layers.

ILLUSTRATION BY Erik Vrielink

1900 to 1945

1952 to 1995

1996 to present

FINE TUNING: Over the past 150 years, technology has improved hearing aid performance. Early conversation tubes [top left] relied entirely on acoustic amplification. The first electric hearing aids [top right] used a carbon diaphragm to amplify sounds. Hearing aids with vacuum tubes [middle left] processed frequencies differently depending on the wearer's type of hearing loss. With transistors came the first hearing aids worn behind the ear [middle right]. Cochlear implants [bottom left] have revolutionized treatment of profound hearing loss. Today's digital hearing aids [bottom right] transform sound waves into digital binary code for processing, and then back to an analog signal for the wearer to hear.

FROM TOP: HEARING AID MUSEUM (2); LES LADBURY/ALAMY

headphones to samples of noisy sentences. The samples were in pairs: first the speech and noise occurring together, and then the same sample after it had been processed by our program running on the deep neural networks. The sentences, which included phrases such as "It's getting cold in here" and "They ate the lemon pie," were cluttered by two types of noise—a steady humming noise and the babble of many people talking at once. The steady noise was similar to the sound of a refrigerator running, in which the audio waves are repetitive and the shape of the frequency spectrum does not change over time. We created the noisy background babble by adding utterances from four male and four female speakers, to mimic a cocktail party.

People in both groups showed a big improvement in their ability to comprehend sentences amid noise after the sentences were processed through our program. People with hearing impairment could decipher only 29 percent of words muddled by babble without the program, but they understood 84 percent after the processing. Several went from understanding only 10 percent of words in the original sample to comprehending around 90 percent with the program. There were similar gains for the steady-noise scenario with hearing-impaired subjects—they went from 36 percent to 82 percent comprehension.

Even people with normal hearing were able to better understand noisy sentences, which means our program could someday help far more people than we originally anticipated. Listeners with normal hearing understood 37 percent of the words spoken amid steady noise without the program and 80 percent with it. For the babble, their performance improved from 42 percent of the words to 78 percent.

One of the most intriguing results of our experiment came when we asked, Could people with hearing impairment who are assisted by our program actually outperform those with normal hearing? Remarkably, the answer is yes. Listeners with hearing impairment who used our program understood nearly 20 percent more words in the babble and about 15 percent more words in steady noise than those with normal hearing who relied solely on their own auditory system to separate speech from noise. With these results, our program built from deep neural networks has come the closest to solving the cocktail party problem of any effort to date.

There are, of course, limits to the program's abilities. For example, in our samples, the type of noise that obscured speech was still quite similar to the type of noise the program had been trained to classify. To function in real life, a program will need to quickly learn to filter out many types of noise, including types different from the ones it has already encountered. For example, the hiss of a ventilation system is different from the hum of a refrigerator compressor. Also, the noisy samples we used did not feature reverberations from the walls and objects in a room, which compounds the noise problem at any cocktail party.

Since we published those early results, we've purchased a database of sound effects designed for filmmakers and used its 10,000 noises to further train the program. Last year, we found that the retrained program could encounter completely new noises and achieve meaningful improvement in comprehension for both hearing-impaired listeners and those with normal hearing. Now, with funding from the National Institute on Deafness and Other Communication Disorders, we are pushing the program to operate in more environments and test it with more listeners who have hearing loss.

Eventually, we believe the program could be trained on powerful computers and embedded directly into a hearing aid, or paired with a smartphone via a wireless link, such as Bluetooth, to feed the processed signal in real time to an earpiece. Periodically, hearing aid wearers could update their devices as manufacturers release new versions after retraining the system on new noises. We have filed several patents for the technique and are working with partners to commercialize it, including Starkey Hearing Technologies, in Eden Prairie, Minn., a leading hearing aid manufacturer in the United States.

With this approach, the cocktail party problem does not look nearly as daunting as it did just a couple of years ago. We, and others, can now create software that we expect will ultimately overcome it through more extensive training in more noisy situations. In fact, I suspect this process is similar to the way children learn to separate speech from noise early in life—through repeated exposure to a wide range of both. With more experience, the approach can only get better. That's the beauty of it. As is also true for a youngster, time is on our side. ■

↗ POST YOUR COMMENTS at http://spectrum.ieee.org/hearingaid0317

Octo-

FLEXIBLE POWER: With no rigid skeleton, the octopus uses sophisticated musculature to move its limbs. Roboticists can mimic its muscles with springs that shorten when a current is applied.

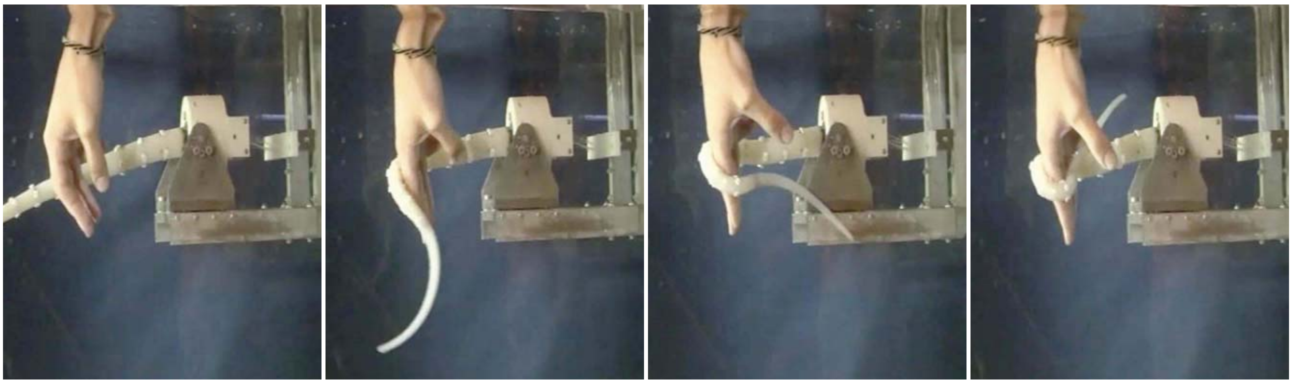
MASSIMO BREGA/THE LIGHTHOUSE (2)

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

A robot octopus points the way to soft robotics

BY CECILIA LASCHI

Bot



The sun was sparkling on the Mediterranean Sea on the afternoon when a graduate student from my lab tossed our prize robot into the water for the first time. I watched nervously as our electronic creation sank beneath the waves.

But the bot didn't falter: When we gave it the command to swim, it filled its expandable mantle with water, then jetted out the fluid to shoot forward. When we ordered it to crawl, it stiffened its eight floppy arms in sequence to push itself along the sandy bottom and over scattered rocks. And when we instructed it to explore a tight space beneath the dock, the robot inserted its soft body into the narrow gap without difficulty.

As a professor at the BioRobotics Institute at the Scuola Superiore Sant'Anna, in Pisa, Italy, I lead a team investigating soft robotics. This relatively new field of research has the potential to upend our ideas about what robots are capable of and where they can be useful. I chose to build robots that mimic the form of the octopus for two reasons. First, because they're well suited to demonstrate the many advantages that come when a machine can flex and squish as needed. Also, it's an excellent engineering challenge: An octopus with eight wiggly arms, which must work together in the face of complex hydrodynamic forces, is very difficult to design and control.

GETTING A GRIP: To be useful, a robot octopus's arms must be able to grasp objects.

In the course of our research, my team hoped to provoke a fundamental rethinking of robotic theories and techniques. We wanted to showcase materials that could be used in actuators that bend and stretch. More crucially, we sought to develop strategies for operating a robot that can curl its limbs in any direction, making it far more tricky to control than a rigid, articulated robot with limbs that have just a few degrees of freedom. To address these challenges, we drew inspiration from nature's design of the remarkable flesh-and-blood octopus.

Up until recently, robots have mainly been used in factories, where their rigid arms are appropriate for the repetitive tasks at hand and the accuracy required. Now, however, roboticists want to put their creations to work in more unpredictable settings where conventional robots often run into trouble.

Some researchers want to build flexible robots that can navigate irregular landscapes, like the ocean floor or the surface of Mars. These robots must move over rough terrain without getting stuck and need manipulators that can grab whatever strangely shaped objects they encounter. Other researchers are focusing on soft robots that can be trusted not to hurt the people they come into contact with. Such soft robots could, for example, work as aides for the disabled or the elderly, and miniature soft robots could even serve as surgical tools inside the body.

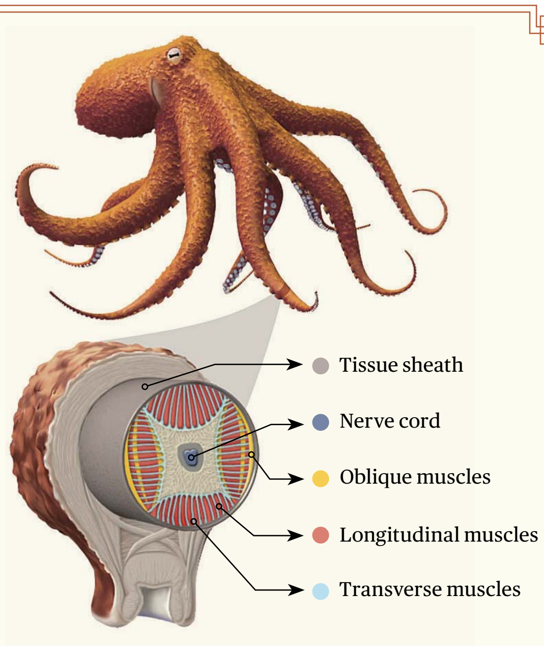
In pursuit of these goals, robotics researchers are increasingly studying animals. That makes sense because

PREVIOUS PAGE, RIGHT: GUSTAV FISCHER/WOODS HOLE OCEANOGRAPHIC INSTITUTION; THIS PAGE: THE BIOROBOTICS INSTITUTE/SCUOLA SUPERIORE SANT'ANNA (4) the bodies of animals are composed mostly of soft materials, with pliable joints and tissue that can change shape without damage. Because their soft tissues absorb shocks and can conform to varied surfaces, animals can use simple control strategies that don't demand great precision.

That, in a nutshell, is why I helped launch the Octopus Integrating Project. The effort brought together several labs from European and Israeli universities, which began collaborating in 2009 to build a robot replica of the fascinating animal. Some of the consortium members had worked on a previous effort that resulted in an "OctArm" attached to a tanklike robot, and they eagerly joined the new effort to copy the animal's remarkable capabilities. We knew it wouldn't be easy.

The octopus has neither an internal nor external skeleton, and its eight arms can bend at any point, elongate and shorten, and stiffen to apply force. It can twist its arms around objects and manipulate them with great dexterity, as demonstrated in plenty of entertaining YouTube videos, including one where the animal steals a camera from an underwater photographer and another where it releases itself from a jar by unscrewing the lid from the inside. An octopus needs such dexterity to survive in the wild. When it crawls along the seafloor, for example, its arms must coordinate their movements in a complex rippling sequence to push and pull its body forward.

Everyone wanted to build a robot that could replicate those agile motions. My team in Pisa, which included my colleague Matteo Cianchetti and others in the BioRobotics Institute, started by studying the octopus arm's muscular hydrostat structure, which allows the overall volume of the arm to remain constant while individual muscles contract and change shape. So when the diameter of an arm decreases, its length increases, and vice versa. To translate biology into engineering, we worked with marine biologists to take measurements of octopus arms and make computer models that could inform our designs. Then we began experimenting with soft actuators that could mimic the animal's muscles.



OCTO-ANATOMY: In a real octopus's arm, a crisscross arrangement of muscles provides all the movement. When the longitudinal muscles contract, the arm gets shorter and fatter; when the transverse muscles contract, the arm gets longer and skinnier.

One option was to make artificial muscles using materials known as electro-active polymers (EAPs). A layer of a soft material is sandwiched between two electrodes; when a voltage is applied, the EAP acts as a capacitor and the electrodes draw closer together, squeezing the soft material between them. Exploiting this phenomenon, researchers have created contractile units that can be arranged in stacks to generate significant forces. A research network in Europe is actively pursuing EAPs for artificial muscles.

Another possibility was to construct our robotic arms using fluidic actuators, in which liquids or gases fill soft chambers to change the shape of the larger structure. Clever design of the shapes and arrangement of the compartments allow a robotic arm to bend in the desired directions and may eventually enable more complicated movements.

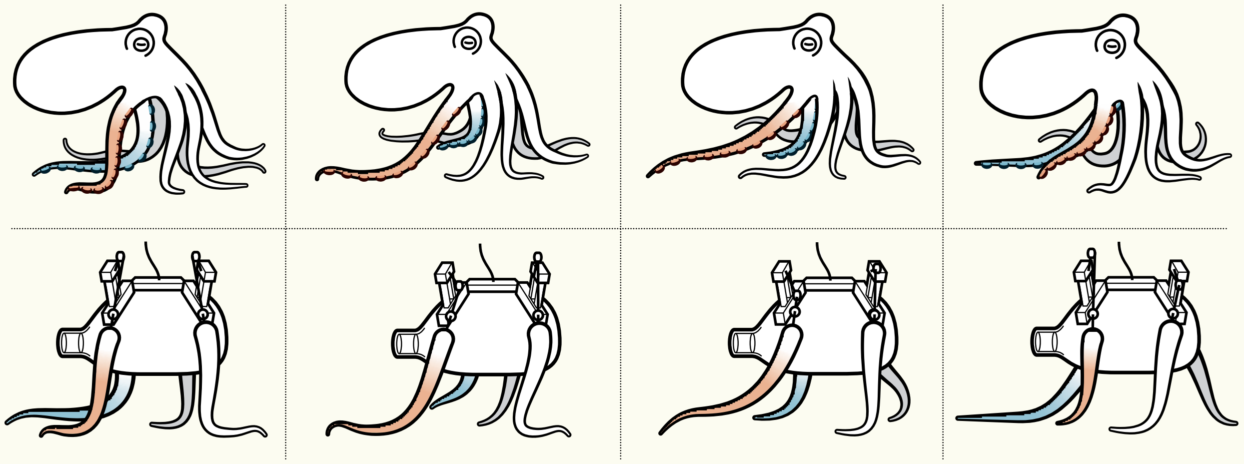
Yet another interesting approach relies on filling a chamber with a granular material, such as sand or even ground coffee, instead of a fluid. With this technique, called jamming, the soft robot remains pliable until a vacuum is applied. Then the robot's body stiffens into a hard shape—like a vacuum-packed brick of coffee on the grocery-store shelf. By applying vacuum to discrete sections in programmed sequences, researchers can make soft robots stiffen and move in specific ways.

My team was most interested in creating artificial muscles using materials called shape-memory alloys (SMAs). When heated, SMAs deform to a predefined shape, which they "remember." We fashioned SMA wires into springs and ran electric current through them to heat them, causing the springs to scrunch up in a way that imitates muscular contractions. For the Octopus project, we constructed a prototype arm using SMA springs to stand in for the longitudinal and transverse muscles found in the limbs of a real octopus. By sending current through different sets of springs, we made the arm bend at multiple points, shorten and elongate, even grasp things.

Our work is primarily meant to demonstrate the potential of soft robotics, and much work remains before a robot octopus will be ready to crawl out of the lab. For example, a bot with sensors on its limbs could provide feedback about its position and the materials it encoun-

ILLUSTRATION BY Emily Cooper

Octopus Locomotion



Attaching [red] *Detaching [blue]*

Elongating [red] *Shortening [blue]*

Detaching [red] *Attaching [blue]*

Shortening [red] *Elongating [blue]*

STEP BY STEP: Studying the octopus's crawling technique reveals a four-stage process, which roboticists can copy to design octo-bots that scoot along the seafloor. First the animal attaches one arm to a firm surface, then it elongates that arm to push its body forward, then it detaches the limb from the surface, and finally the octopus shortens the arm so it catches up with the body.

ters, which could lead to better control strategies. A team of researchers at the Worcester Polytechnic Institute, in Massachusetts, is addressing just that challenge by embedding proprioceptive sensors in a robotic snake.

It's fun to imagine how an advanced robot octopus with eight dexterous arms could perform in the wild. Take the marine-energy industry, where there's interest in placing tidal turbines on the seabed to harvest power from flowing water. If the machinery breaks, repairs would be difficult and expensive: Workers would have to haul turbines up to the surface or send human divers down. Maybe, one day, an octo-bot technician could be sent instead. With its agile limbs, it could manipulate tools and make the fix.

We roboticists aren't interested in the octopus for its limbs and muscles alone—we also value its particular brand of intelligence. The octopus's brain and peripheral nervous system are well developed compared with those of other mollusks, but they're still fairly limited. It's surprising, then, that they can control a huge range of movements in eight independent arms. So our next challenge under the Octopus project was to study the animal's control mechanisms, looking for ways to manage a flexible robot's complex movements.

Biologists have determined that the octopus's brain doesn't issue top-down commands for every small movement of its twisty limbs. In *Octopus vulgaris*, the common octopus, the brain actually contains far fewer neurons than the peripheral nervous system. Biologists believe that the brain initiates motions, while lower motor centers control the precise neuromuscular activity. Experiments have shown that even if you sever the nerves descending from an octopus's brain, its arms can still recoil from unpleasant stimuli and reach out as if to grab something.

And here's what we found even more interesting: The octopus's limbs don't need comprehensive directions to produce the desired movement. Thanks to millions of years of evolution, their bodies are designed to respond to their environment in certain automatic and useful ways. This concept is often called morphological computation by roboticists, while artificial intelligence researchers refer to it as embodied intelligence.

When translated to the robot world, this principle means we should design our robots so that the physical properties of their bodies automatically produce the desired movements. With this strategy, extremely simple commands can cause a robot to efficiently carry out complex tasks.

My team, including my colleague Marcello Calisti and others, kept this principle in mind when we first set out to make a robot octopus that could crawl along the seafloor. We studied the locomotion strategy of the real ani-

ILLUSTRATION BY James Provost

mal and determined that it uses a four-step procedure to crawl. First it attaches one of its rear limbs, which are covered in suckers, to the bottom. Then it elongates that limb to push the rest of its body forward. In the third stage it pulls up its suckers to detach the limb from the seafloor. And finally, it shortens the limb to bring it closer to its body and prepare for the next push. The animal ripples forward as two of its arms perform these actions in graceful sequence.

To mimic this form of locomotion, it might seem that our octo-bot would need to carry out a huge number of computations to control its eight arms, each of which can bend at any point. In a typical robot, each degree of freedom for a limb requires an actuator to drive movement in that direction and some sort of numerical controller to govern the actuator. But our robotic arms had too many degrees of freedom to apply the usual control strategy.

We decided to follow the example of evolution and try to fashion arms that didn't require complex control inputs. Before we built anything, we formulated mathematical models to test various aspects of the arm's design, including the density of the material used, its stiffness, its shape, the internal placement of its artificial muscles, and so on. All of these parameters needed to be examined in the context of how the arms would perform underwater. How would these arms fare in water of different salinities and temperatures? How would increased depth and pressure affect them? How would currents and turbulence influence their movement? To answer those questions, we added hydrodynamic factors to our models. We also had to consider the texture and composition of the surfaces over which they'd be crawling. We decided not to complicate our design by adding suckers to the arms, instead using a material for the exterior that would produce plenty of friction.

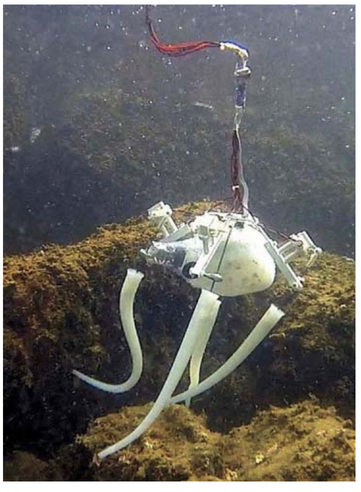
With all those parameters in the model, the task of finding the optimal combination became far too complex to compute by trial and error, so we used an evolutionary algorithm to explore the vast range of possibilities. This algorithm started by creating many hypothetical octopus body shapes, each with its own set of characteristics. Then it tested those octopuses' arms to see which performed best in the models, and used the "fittest" limbs' attributes to inspire a new batch of possibilities. In this way, we identified an arrangement that would generate the correct amount of propulsive force and produce the desired crawling movement.

The results were very gratifying: It proved simple to control the octo-bot, and we could mimic the four-step crawling we'd observed in the real animal. For this prototype, which was intended to showcase control mechanisms rather than materials, we swapped out the SMAs and instead used cables as artificial muscles. Each silicone-rubber arm contained a steel cable that lengthened or shortened the arm and a carbon-fiber cable that bent the arm for the attachment and detachment steps (and also enabled the robot to grasp objects). One simple servomotor in each arm provided power.

An enormous number of sophisticated computations had been done at the time of the design. But when the robot was in action, we said proudly that it was very dumb: It had only a simple "brain," or microcontroller, which triggered the arms in the correct sequence to make the octo-bot crawl by virtue of the mechanical properties of its body. Our octo-bot demonstrated morphological computation at work.

In 2012, my team at the BioRobotics Institute began a related project called PoseiDrone. For that venture, we aimed to create a soft underwater robot that could not only crawl and grasp objects but also swim.

To give this robot swimming abilities, we once more relied on morphological computing. The key was the design of the octo-bot's mantle—the headlike part of an octopus that swells up when it fills with water, then contracts to eject the water and generate jet propulsion. Again, we used computer models to determine the size, shape, and material properties of the silicone mantle, paying careful attention to the ways the structure would deform as it ejected the water. Our algorithm produced the optimal combination that would, with only one small motor and a few simple cables, send the octo-bot jetting through the water.

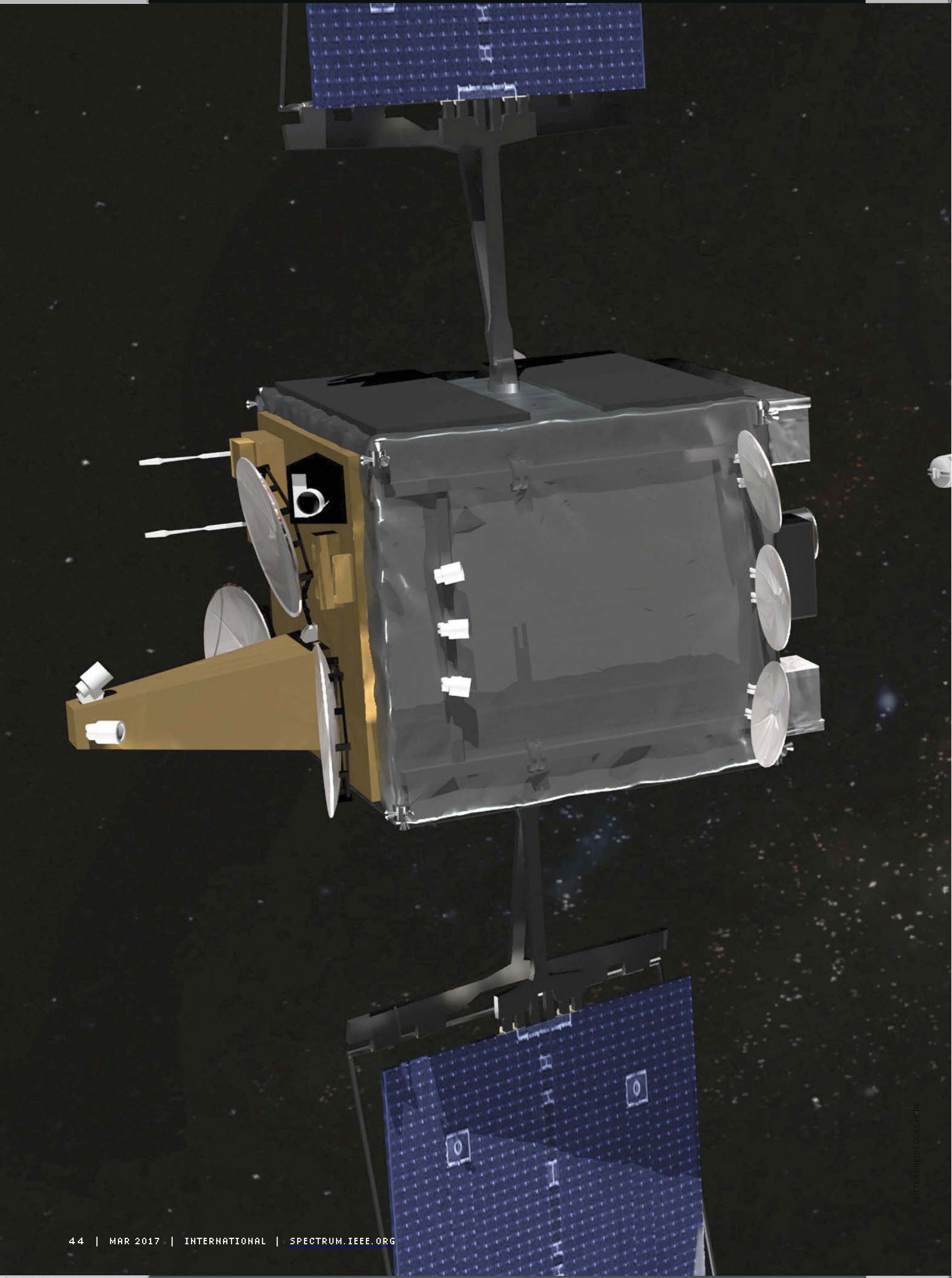


SWIMMING LESSONS: *A robot imitates the octopus's jet propulsion method to move through Mediterranean waters.*

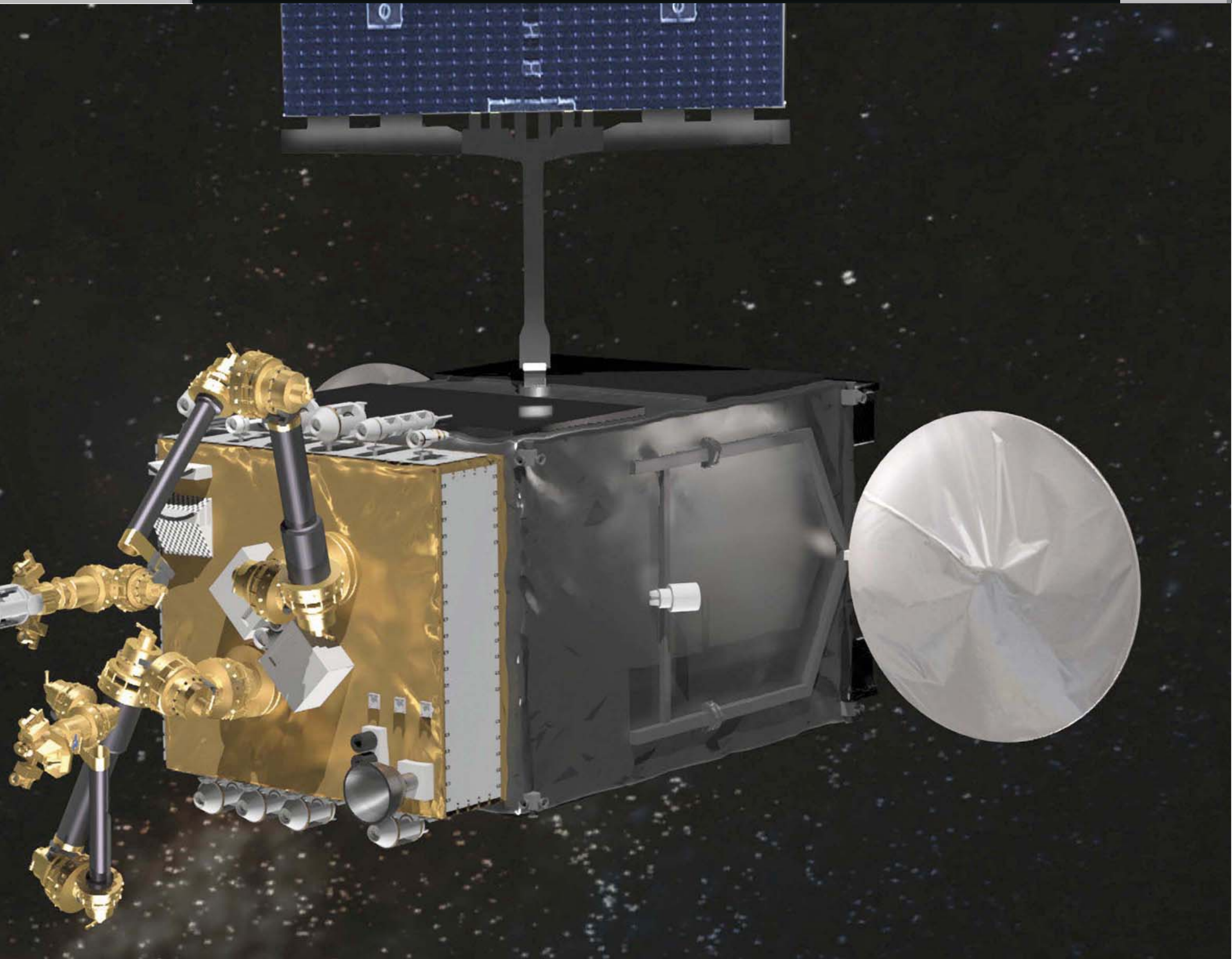
MASSIMO BREGA/THE LIGHTHOUSE

With our PoseiDrone prototype, we studied its ability to perform various tasks, such as crawling while toting an object and propelling itself with a series of bursts while swimming. That prototype was the bot we tossed into the Mediterranean to see how it would fare in the great outdoors. We were delighted to watch our robot deal handily with the unpredictable surfaces, waves, and currents it encountered, and at the conclusion of our experiments we found a suitable retirement for this exemplary specimen. The PoseiDrone robot will soon reside in a tank in the Livorno, Italy, aquarium—right next to a real octopus. ■

↗ POST YOUR COMMENTS at http://spectrum.ieee.org/octobot0317



IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page



ORBITAL MECHANICS

Soon, robots will repair and refuel satellites in space

By **Gordon Roesler, Paul Jaffe & Glen Henshaw**

LET'S SAY YOU ARE THE PROGRAM MANAGER of a very large, complex system. Perhaps it's an aircraft, or a building, or a communications network. Your system is valued at over US $500 million. Could you imagine being told that you won't ever be able to maintain it? That, once it's operational, it will never be inspected, repaired, or upgraded with new hardware? ● Welcome to the world of satellite building. After a satellite is launched, it is on a one-way journey to disrepair and obsolescence, and there is little anyone can do to alter that path. Faults (which are called anomalies in the space business) can only be diagnosed remotely, using data and inferential reasoning. Software fixes and upgrades may be possible, but most mechanical anomalies can't be corrected. And hardware can't be replaced, which means that even if a satellite is operating well, it could lose its state-of-the-art status just a few years into a typical 15-year lifetime. ● This is a lost opportunity. At this moment, more than 300 commercial satellites are providing television, radio, data transfer, weather, and telephony from geosynchronous earth orbit (GEO),

IEEE SPECTRUM Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

some 36,000 kilometers above the Earth's surface. Dozens of government satellites are there too. Each of those satellites costs hundreds of millions of dollars to build and launch. If governments and private companies could actively repair and revitalize their satellites in GEO–and move them to new orbits as needed–they could extend their investments and substantially defer the cost of building and launching replacements.

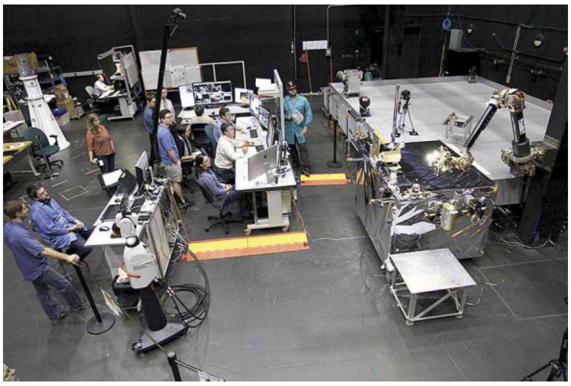
To be sure, people have already done such "servicing" work. In the past, astronauts have captured orbiting satellites–most notably in a series of space shuttle missions to repair and upgrade the Hubble Space Telescope. But these operations have been performed only in low Earth orbit, many thousands of kilometers closer to Earth, and each mission puts human lives on the line. Sending humans out to GEO would put them at even further risk, as they'd be exposed to more intense radiation from solar particles and cosmic rays.

The bottom line is this: If we hope to make satellite repair economical, safe, and routine, we'll have to turn to robotics–that is, build spacecraft with dexterous robotic arms and design them so they can perform a range of delicate operations, often on their own.

The idea of building such robotic satellite servicers isn't new, but we think it is finally starting to get some traction. We are now working on a project sponsored by the U.S. Defense Advanced Research Projects Agency (DARPA) to develop a robotic servicing spacecraft that can work on satellites that were never designed to be serviced–which is pretty much all of them in orbit today. This public-private partnership, the Robotic Servicing of Geosynchronous Satellites (RSGS) program, builds on a decade of work by DARPA (where coauthor Roesler works) and the U.S. Naval Research Laboratory (where Jaffe and Henshaw work), as well as the efforts of university researchers and space agencies around the world.

When RSGS launches in the early 2020s, we will be able to work with satellite owners to move GEO satellites to new orbits, fix stuck solar panels, and perform several other important chores. Independently, NASA plans to launch a robotic mission around the same time called Restore-L, which aims to refuel and relocate a government-owned satellite in low Earth orbit.

If successful, these two missions will push the limits of automation and robotic operation in space. They could be the first steps toward space construction projects like vast solar arrays that can beam energy back to Earth, robots that could mine asteroids and deflect those that pose a danger to Earth, and many other applications that would revolutionize the way we operate in space and what we can accomplish there.



A JOINT EFFORT: RSGS will require a good deal of autonomy in order to safely work with other spacecraft in geosynchronous orbit. But a human will always be in the loop.

THE DEVELOPMENT OF ROBOTIC SERVICING has been stymied by a classic chicken-and-egg problem: Satellites aren't designed to be serviced because there are no robotic servicers out there they need to be compatible with, and no one builds robotic servicers because there are no satellites for them to service.

This situation has persisted for decades, in large part because servicing satellites robotically is a really difficult problem. Satellites are built to be as lightweight as possible to keep launch costs down, and their systems are tightly integrated for the same reason. Even opening up a completed spacecraft to replace a part when it's still in the factory is a major proposition.

To do even the most basic operation on a satellite, a robotic repair vehicle must be able to manipulate a very expensive, relatively fragile device under challenging conditions. Among these are high radiation levels, wildly varying temperatures and illumination, and time delays that dramatically complicate remote control. To add to the difficulty, an inadvertent tap can send something spinning away through the drag-free, frictionless environment.

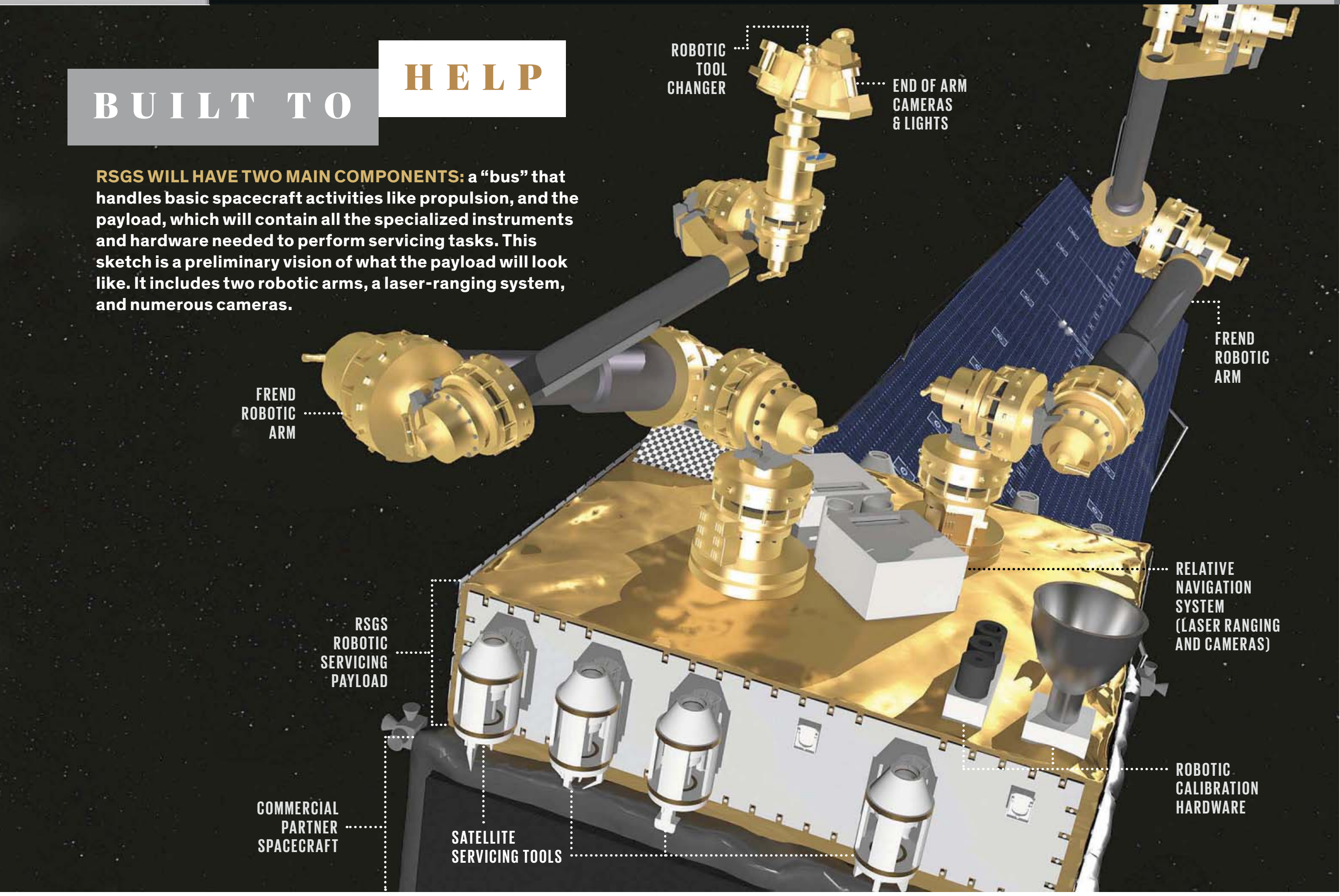
To be successful, a robotic servicer's electronic and mechanical systems must operate reliably because a computer crash or a motor failure could damage the patient. The robot must carry enough onboard intelligence to know how to avoid unintentionally hitting itself with its own arm or slamming into the client spacecraft. Both the position of its arms and the force they exert must be controlled with great precision. Above all it needs to be able to understand when things are going well, when they aren't, and what to do in response.

Even in a terrestrial laboratory, these are challenging criteria, but roboticists have made great strides toward meeting them. The first in-space demonstration was done in the late 1990s by a Japanese flight experiment called Engineering Test Satellite VII, or ETS-VII. In 2007, the United States conducted a more ambitious on-orbit experiment, the DARPA Orbital Express program. ETS-VII had shown that a service robot can autonomously rendezvous and dock with another satellite. In addition to those tasks, Orbital Express demonstrated both fuel and component transfer. But neither ETS-VII nor Orbital Express solved the chicken-and-egg problem. Each mission launched two spacecraft–a servicer and a client satellite–and each client was specially designed to be compatible with the servicer's robotic hardware.

The demonstrations were useful, but they weren't sufficient to convince satellite makers to alter their designs to make them compatible with robotic maintenance (and of course, existing satellites would be ineligible for help). So in 2002, DARPA set out to answer a different question: What can a robotic ser-

BUILT TO HELP

RSGS WILL HAVE TWO MAIN COMPONENTS: **a "bus" that handles basic spacecraft activities like propulsion, and the payload, which will contain all the specialized instruments and hardware needed to perform servicing tasks. This sketch is a preliminary vision of what the payload will look like. It includes two robotic arms, a laser-ranging system, and numerous cameras.**



vicer do for the sorts of satellites we have today, which aren't made to be serviced? To help find the answer, the agency sponsored studies by researchers at NASA's Jet Propulsion Laboratory, in Pasadena, Calif., and at the Naval Center for Space Technology (NCST) of the U.S. Naval Research Laboratory in Washington, D.C.–the oldest continuously operating satellite design and testing facility in the United States.

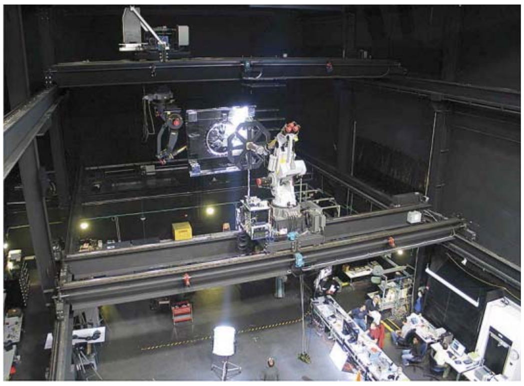
One study, called RescueSat, considered a fairly common scenario, in which a satellite reaches space but a rocket or motor failure prevents it from getting to its intended orbit. RescueSat concluded that at least one problem with servicing could be solved: the process of grabbing hold of the satellite, or "grappling." It turns out that every satellite has structures that connect it to its booster, which separates shortly after launch. What remains after separation are large aluminum rings or bolt holes, which are exposed and strong. These can be grappled fairly easily by the robotic arms of a servicer. This robotic embrace would turn the two independently moving satellites into one rigid two-satellite assembly that could remain steady during servicing.

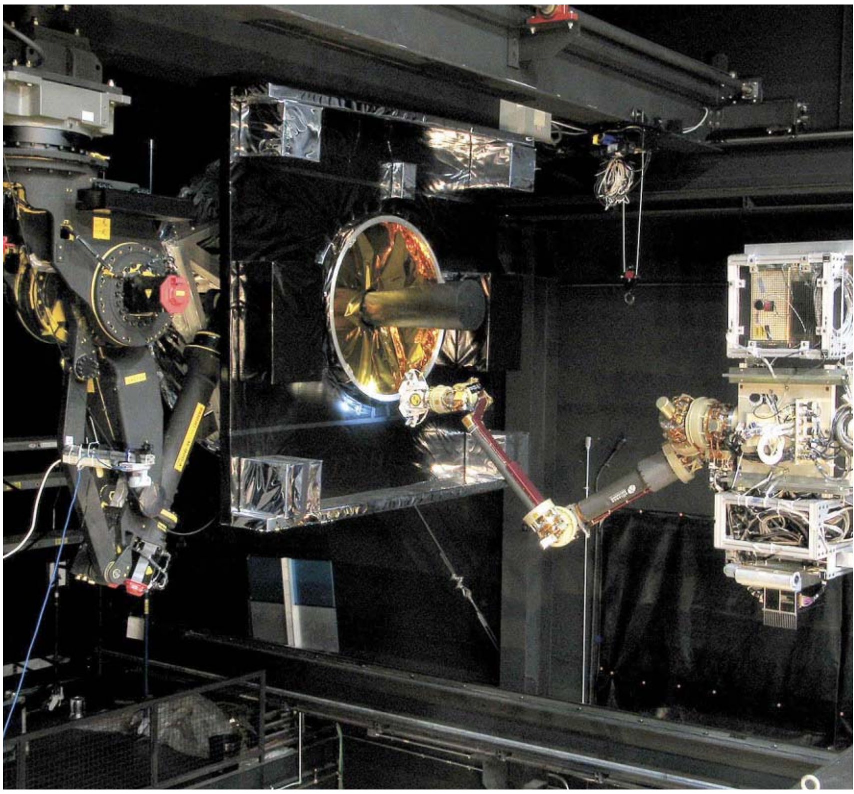
If a servicer can grapple a satellite with these exposed booster-attachment features, it should be able to push the satellite around and move it to a new orbit. But "should be" isn't conclusive enough to convince satellite owners to allow a servicer to approach their spacecraft. So RescueSat was followed by a second study called Spacecraft for the Universal Modification of Orbits (SUMO), which tested the approach with real hardware, including a pair of commercially available robotic arms. This trial took advantage of a newly constructed facility at NCST that features a computer-controlled crane system that can move test satellites around as if they were in orbit. This system allows engineers to simulate rendezvous operations with realistic orbital mechanics, where firing a thruster in order to speed up or slow down will also put your spacecraft in a new orbit.

The main focus of the SUMO tests was autonomy, which is arguably the biggest technical challenge to robotic servicing. Owing largely to data-processing delays, it can take several seconds for pictures from a satellite in GEO to reach human operators on the ground, and the same interval again for a command to get back up to the satellite. This delay makes teleoperation quite challenging and even dangerous, so at certain times, such as during the last few meters of a servicer's approach to a satellite, the servicer needs to be able to control itself–and be able to abort if anything seems to be going wrong.

By 2005, SUMO showed that it should be possible for a robotic servicer to safely and reliably approach and grab hold of another satellite using automated software. The project was a good proof of principle, but to get a more realistic simulation, we would need to test flight-quality robotic arms, ones that would be sturdy enough to handle the docking of two multiton objects in orbit. So in 2005, DARPA selected California-based Alliance Spacesystems, which built the robotic arm on NASA's Curiosity rover, to construct a flight prototype under a program called Front-end Robotics Enabling Near-term Demonstration (FREND).

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY





ARM TESTS: **A crane system [top left] is used for RSGS tests at the U.S. Naval Research Laboratory. A human operator can monitor and plan actions [bottom left], such as approach and grapple [right] of a mock satellite.**

U.S. NAVAL RESEARCH LABORATORY (3)

Completed in 2008, the FREND arm was able to pass the rigorous tests required of hardware to fly in space. And it enabled the team at NCST to conduct more realistic experiments, using the simulator facility to test rendezvous-and-grappling processes under conditions that mimicked those in space.

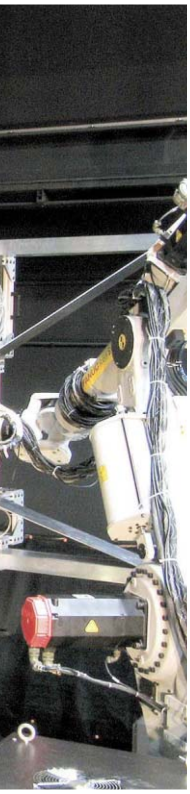
One difficult factor that had to be simulated and addressed was lighting. Many parts of satellites are covered with thermal blankets that can be made of such materials as aluminized Mylar, the same material used in rescue blankets. Illuminated by the sunlight, some spacecraft parts may "sparkle." Machine vision is used to process camera images, but the highly reflective, high-dynamic-range conditions in space are quite challenging, and they can cause the spacecraft to over- or underestimate distance. Rather than trying to solve this problem in software, laser-ranging sensors can be used to overcome this effect.

Another challenge was processing power. Space-qualified computers typically lag the state of the art by several generations. But during the SUMO and FREND programs, we demonstrated that we could implement all the needed autonomy, machine vision, and control software on CPUs considerably slower than those used in current terrestrial robots, and that we could make these slower systems work together in a reliable way.

You may have heard about satellite servicing a few years back. After FREND was complete, DARPA began a program called Phoenix, which looked at the possibility of harvesting large antennas from "dead" satellites and giving them new life by attaching electronics and propulsive modules. This very appealing idea would have been a first step toward doing something useful with what was otherwise "space junk." But ultimately, DARPA decided that servicing active satellites in GEO would result in a higher payoff than refurbishing defunct spacecraft. Still, the tools and techniques developed in Phoenix are now available for the RSGS mission, which aims to offer commercial and military satellite operators a wide range of services to extend the lives of existing satellites.

EVERYTHING WE'VE LEARNED from previous projects will go into RSGS. A commercial provider will build the spacecraft "bus"—the base of the spacecraft that will supply power and handle communications with the ground. The Naval Research Laboratory will assemble and test the payload to go atop it. This payload will consist of a laser-ranging system to help with long-distance targeting, a dozen cameras, two robotic arms with cameras mounted on the ends, more than 100 circuit boards, and a tool kit of appendages the arms can swap in and out like a multibit screwdriver. The arms—each 88 kilograms and 2.3 meters long—are improved versions of the FREND arm. The FREND team in Pasadena is now a subsidiary of MacDonald, Dettwiler and Associates, which built robotic arms for the space shuttles and the International Space Station.

RSGS will aim for geosynchronous orbit, and because GEO satellites orbit in or near a single plane, the RSGS servicer will be able to visit dozens of them throughout its planned 5- to 8-year lifetime.



The spacecraft won't be able to open up those tightly packaged satellites to swap out fried circuit boards and reseat connectors. But it could help correct some common mechanical problems, by using one of its robotic arms to hold onto a satellite while the other gently prods and "unsticks" a solar panel or antenna. These appendages are folded up for launch and are meant to unfold after they reach space. But once every two or three years, an antenna or solar panel does not deploy properly—and a mission costing hundreds of millions of dollars comes to naught.

These will be delicate operations. Antennas and solar panels are designed to be as lightweight as possible; in fact, spacecraft solar panel assemblies aren't even strong enough to support their own weight in gravity. So, in order to ensure the robot is gentle enough to manipulate client spacecraft components without damaging them, we implemented a technique called compliance control. The force applied by an arm is constantly measured, and the robot joints are commanded to respond appropriately so the applied force always stays within a safe margin.

This sort of feedback will be important in space. Wide temperature swings alter dozens of parameters on the arm, including stiffness and coefficients of friction. In the current design, the RSGS arms will be able to push on a small spring mounted on the exterior of the servicer to make sure their force and torque sensors are calibrated properly throughout the mission.

Another valuable service that we began exploring with SUMO is pushing client satellites to new orbits. This could be done to help deliver a satellite to its intended orbit if it has failed to get there or to modify an existing orbit to save the spacecraft's onboard fuel. RSGS's repositioning capacity could also be useful at the end of a satellite's life. Retiring satellites are moved some 300 km above GEO to ensure that the valuable orbit doesn't get filled up with nonoperational satellites that could collide with others and create space debris. A servicer could be used to ferry satellites to this outer "graveyard" orbit, which would extend the satellites' lives by allowing them to use the last of their fuel for routine operations.

A particularly exciting capability being developed for RSGS is to attach small new payloads to operating satellites. These modules would have to carry their own power and communications gear, but they could piggyback on an operating satellite's positioning and propulsion capabilities to stay in the correct orbit and pointed in the right direction.

Because of communications delays—and the time it would take for an earthbound human operator to analyze the spacecraft's position and other factors—it would be impractical and potentially dangerous to try to conduct these sorts of activities by remote control. As a result, we expect RSGS will use a mix of human and autonomous operation called supervised autonomy. A human will always be in the loop, monitoring RSGS's activities, but for certain critical tasks such as grappling, the spacecraft will be on its own, performing either a prescripted activity or operating in a fully autonomous mode by which the servicer can rapidly respond if something does not go according to plan.

RSGS engineers are still working on this capability and expect to conduct more tests over the next few years. That effort includes the onboard software needed to control the servicer, which is expected to run to about a million lines of code—not unheard of for a spacecraft, but definitely more than the average satellite.

SPACE ROBOTICS CAPABILITIES that are being developed for RSGS and NASA's Restore-L could be extended, in the long term, well beyond servicing existing satellites. Today, great effort is expended in complicated (and sometimes unreliable) mechanisms to maximize the size of mirrors and antennas that can be deployed from a single launch with a single satellite. But a dexterous and capable space robot could assemble pieces from multiple launches in space to construct antennas and other structures of enormous size.

On-orbit robotic construction could boost the sensitivity of radar imaging and speed of communications by allowing for much bigger and more sensitive radio antennas. It could also be a boon for Earth-observing satellites and astronomical telescopes.

We could even help address Earth's growing energy demands. Nearly all spacecraft gather energy from solar panels. If we can increase the area of solar panels a thousand- or a millionfold, and send the energy to Earth via microwaves, we could create a new source of clean, constant, and globally transmissible energy [for more, see "It's Always Sunny in Space," *IEEE Spectrum*, May 2014]. This idea is starting to look more economically viable, thanks to the recent success of SpaceX and Blue Origin in recovering expensive booster rockets for reuse, and the ongoing development of modular and mass-producible space components.

Further in the future, robotic construction capabilities could also be applied to create space infrastructure—standing structures such as refueling depots, in-space manufacturing facilities, space-tourism complexes, and asteroid mining stations. And the same sorts of robots that can move satellites could also help keep the space environment around Earth safe, by helping eliminate space debris. Someday, work in this area might even lead to spacecraft capable of diverting asteroids from likely collision courses with Earth.

It may be decades before we can see scores of robots working in space. But within five years, we expect to have systems in orbit that can demonstrate the autonomy, dexterity, and delicacy needed to begin work on these more far-reaching capabilities. Delivering on their promise will require overcoming many technical as well as economic hurdles. But if RSGS succeeds in its mission to demonstrate autonomous servicing in GEO, it will send a message to satellite owners: A new robotic revolution awaits on the high frontier. ■

This article represents the authors' personal views and not necessarily those of DARPA, NRL, or the U.S. Department of Defense.

↗ POST YOUR COMMENTS at http://spectrum.ieee.org/spacerobotics0317

FACULTY POSITION OPEN

Toyota Technological Institute has an opening for a tenured- or tenure-track faculty position in the Department of Advanced Science and Technology. Applications are encouraged from all relevant areas. For more information, please refer to the website **http://www.toyota-ti.ac.jp/english/employment/associate.html**

Position: Associate Professor (tenured- or tenure-track)

Research field: Innovative information storage technology using materials, such as semiconductor, dielectric and magnetic, the related devices, and also systems, for example, spin electronics materials, devices, and systems

Qualifications: A Ph.D. in a relevant field. The successful candidate is expected to demonstrate potential to develop strong and outstanding programs in the above research field. It is also necessary for him/her to supervise students, and to teach advanced and basic courses both at the undergraduate and graduate levels.

Starting date: April 1, 2018 or at the earliest convenience

Documents: (1) A curriculum vitae
(2) A list of publications
(3) Copies of 5 selected papers
(4) Brief description of research activities and future plan for research and education (3 pages each)
(5) Names of two references with e-mail address and phone numbers
(6) Application form available from our website

Candidacy will not be considered unless the full documents are submitted. The application documents will not be returned.

Deadline: April 28, 2017

Inquiry: Search Committee Chair Professor Yasutake Ohishi
Phone: +81-52-809-1860 E-mail: y-ohishi1@toyota-ti.ac.jp

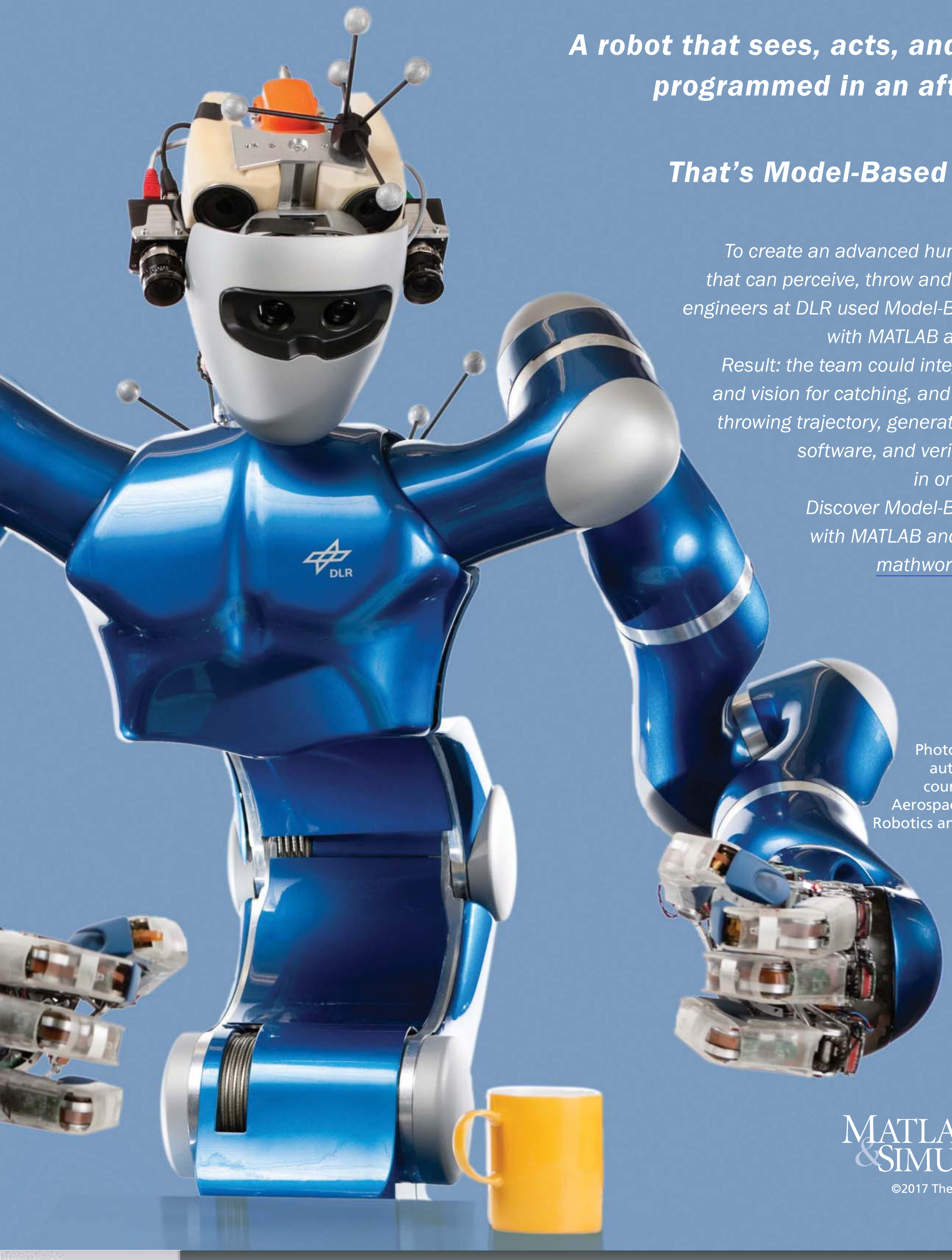
The above should be sent to:
Mr. Masashi Hisamoto
Administration Division
Toyota Technological Institute
2-12-1, Hisakata, Tempaku-ku

(Please be advised to write “Application for Memory Engineering Laboratory” in red on envelope)

WHEN FORM EXCEEDED FUNCTION

In 1960, when Goodyear came out with prototypes of these translucent illuminated tires, it suggested that the tires' color could be changed to match whatever the passengers of the car happened to be wearing. That feature alone wasn't quite enough to justify producing the tires, which didn't perform well in the rain and melted under heavy braking. Today, Goodyear does make tires that match the color of your outfit, so long as you always wear black. ■

GETTY IMAGES


DLR