Is Hydrogen the Next Jet Fuel? › H_2-powered aircraft reach the runway
P.5

These Semiconductors Are Surprisingly Green › They save power to cut carbon
P.32

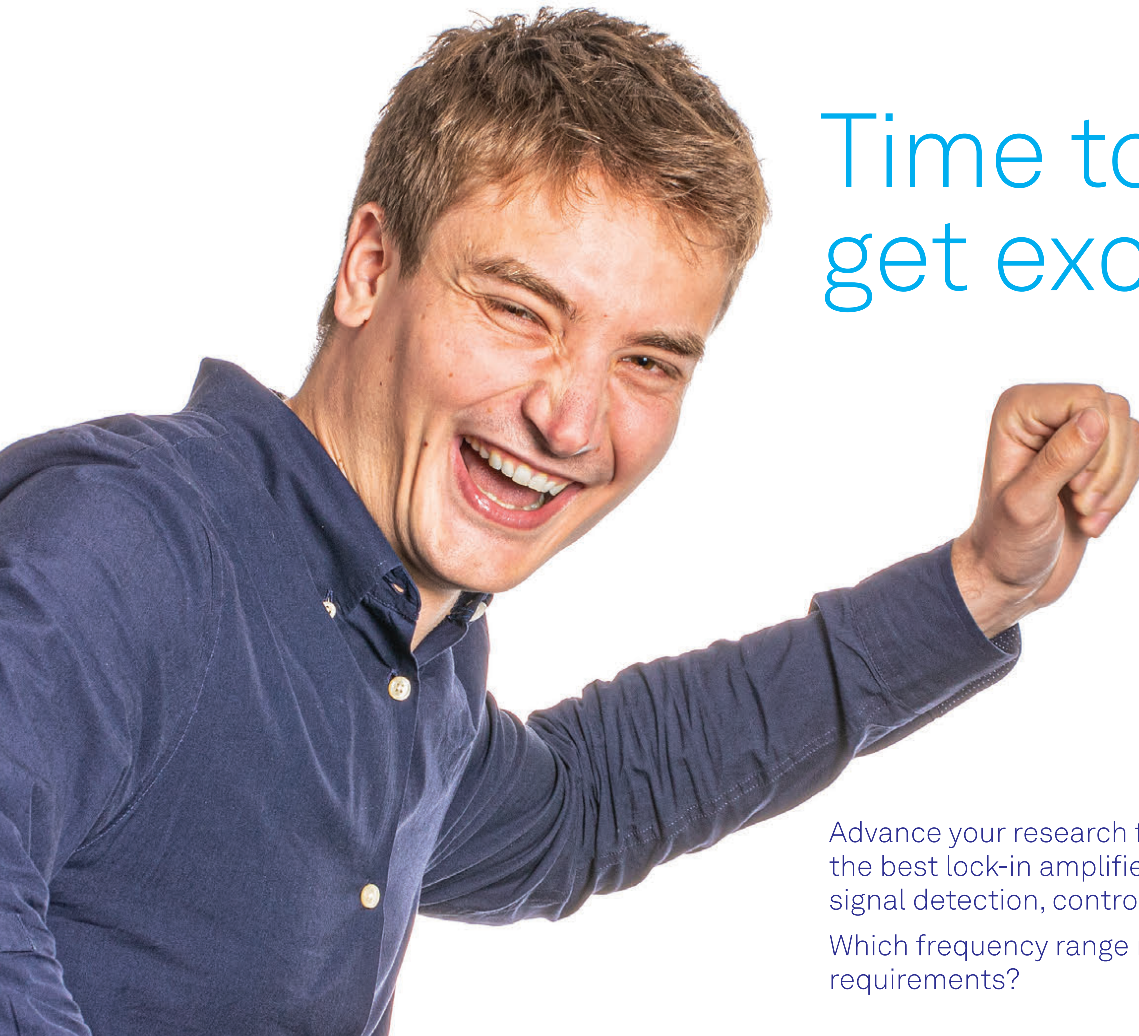
Straight Talk on the EV Transition › What it will really take to electrify road transport
P.40

FOR THE TECHNOLOGY INSIDER
APRIL 2023

IEEE Spectrum

Finding Somerton Man

How DNA and AI facial reconstruction cracked a cold case



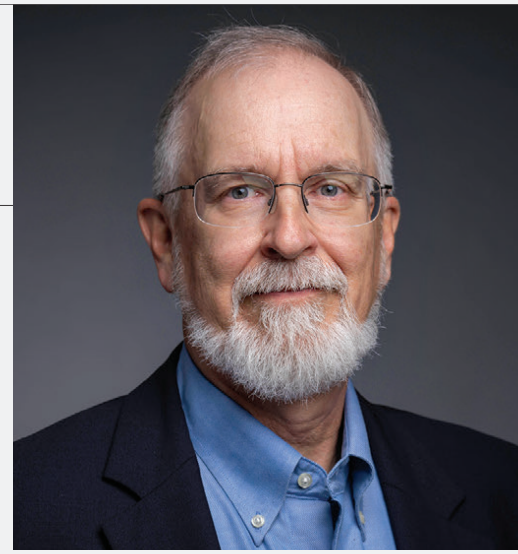
VOLUME 60 / ISSUE 4 APRIL 2023

Contents

ON THE COVER:
Illustration by Mike McQuade

Illustration by Sergio Albiac

BY HARRY GOLDSTEIN



The Staggering Scale of the EV Transition

One engineer grapples with the challenges ahead

Over the last 20 or so years, Contributing Editor Robert N. "Bob" Charette has written about some of the thorniest issues facing the planet at large and engineers in particular. For *IEEE Spectrum*, he's dug into software reliability and maintenance, the so-called STEM crisis, and the automation paradox, examining those complex topics through the eyes of a seasoned risk analyst.

I've been fortunate to be Bob's editor for many of his ambitious projects. We often converse on Friday afternoons about what he's hearing from industry insiders and academics on whatever subject he's currently investigating. Our conversations are jovial, sometimes alarming, and always edifying, at least for me.

So when he called me on a Friday afternoon in the summer of 2021 to propose an article delving into the complexities of the global transition to electric vehicles, I knew that he'd do the research at a deeper level than any tech journalist, and that he'd explore angles that wouldn't even occur to them.

Take power-grid transformers. These essential voltage-converting components are designed to cool down at night, when power consumption is typically low. But with more people charging their EVs at home at night, the 30-year design life of a transformer will drop—to perhaps no more than three years once mass adoption of EVs takes hold. Bob examined factors like that and dozens of others during the last year and a half.

Throughout his reporting, Bob focused on the EV transition "at scale": What needs to happen in order for electric vehicles to displace internal-combustion-engine vehicles and have a measurable impact on climate change by midcentury? Quite a lot, it turns out. Humans must change two foundational sectors of modern civilization—energy and transportation—to achieve the targeted reductions in greenhouse gas emissions. These simultaneous global overhauls will involve trillions of dollars in investments, tens of millions of workers, millions of new EVs, tens of thousands of kilometers of new transmission lines to carry electricity from countless new wind and solar farms, and dozens of new battery plants and new mines to feed them. Then there are the lifestyle compromises that most people living in developed countries will have to make.

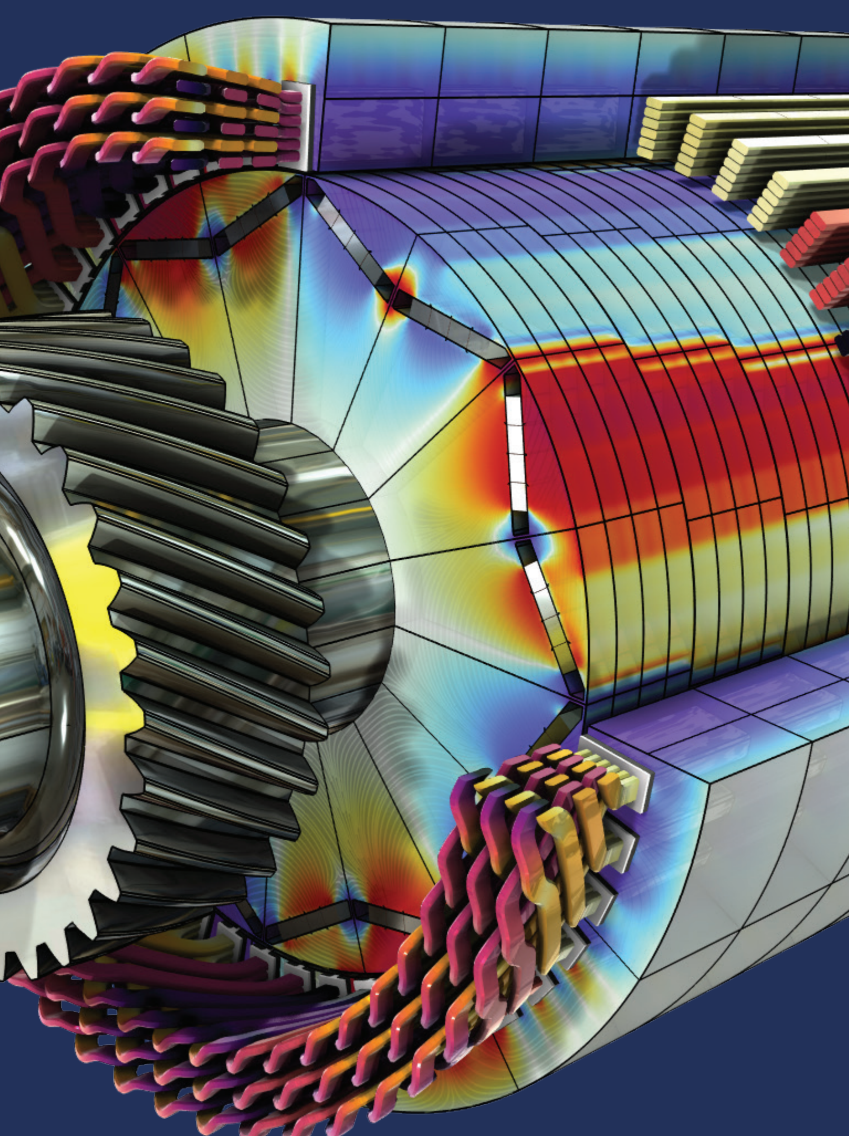
"As always, Bob approached this tangle of issues from a systems engineering perspective."

As always, Bob approached this tangle of issues from a systems engineering perspective. For a typical assignment, his first draft runs to thousands of words over the assigned length, and we manage to chisel the manuscript down to a single feature article, often leaving a lot of interesting material on the cutting-room floor. This time, we decided to let Bob go long, so he could paint as detailed a picture as possible of a fast-moving and multifaceted target.

The process was kind of like building an EV while driving it. As Bob's editor, my job was to help him synthesize his findings into a snapshot of the EV industry at a pivotal moment in history. Every Friday, we'd discuss some new announcement or jaw-dropping data point that needed to find its way into what I came to call The Opus. As Bob talked to more people and read more policy documents, research reports, and public-meeting minutes, the assignment went from two or three articles to twelve, covering the technological hurdles, policy battles, and consumer attitudes surrounding the EV transition.

The Opus is now an e-book, *The EV Transition Explained*, the introduction to which you can read on page 40. The e-book itself is available for download exclusively for IEEE members via our website or the QR code at the end of the article. Bob's hope, and ours, is that policymakers, auto-industry executives, engineers, and consumers will use his analysis to inform their discussions and decisions about the best ways to transition to EVs—at scale. ■

PORTRAIT BY SERGIO ALBIAC; BOB CHARETTE



CONTRIBUTORS

DEREK ABBOTT

Abbott, a fellow of the IEEE, earned his doctorate in electrical and electronics engineering at the University of Adelaide, in Australia, where he is a professor. He likes to use scientific methods to solve longstanding mysteries. He is most famous for identifying a body found on an Australian beach in 1948. He describes how he cracked this old, cold case in "Finding Somerton Man," on page 22.

ROBERT N. CHARETTE

Charette, a senior member of the IEEE, is a contributing editor to *IEEE Spectrum* and an acknowledged international authority on IT and systems risk management. A self-described "risk ecologist," he examines the intersections of business, political, technological, and societal risks. The transition to electric vehicles brings all these intersections into sharp relief. Charette introduces his online series and e-book, *The EV Transition Explained*, on page 40.

MATTHEW N. EISLER

Eisler is a historian of science and technology at the University of Strathclyde, in Scotland, where he studies the relationship between energy and environmental policy and science and engineering practice. In this issue, he relates the tricky history behind efforts to use electric vehicles as resources for the power grid [p. 46]. EVs are the topic of Eisler's second book, *Age of Auto Electric: Environment, Energy, and the Quest for the Sustainable Car* (MIT Press), which came out last year.

UMESH MISHRA

Mishra, an IEEE Fellow, is a professor of electrical and computer engineering at the University of California, Santa Barbara. He is also cofounder and chief technology officer of the gallium-nitride electronics company Transphorm. After earning a Ph.D. at Cornell for research on gallium arsenide–based microwave devices, he had an epiphany: "I realized that the fundamental physics of microwave devices could be applied to power devices, and that the impact on the world, and climate, would be much bigger." On page 32, he explains how gallium nitride and silicon carbide are fulfilling that promise.

EDITORIAL CORRESPONDENCE
IEEE Spectrum, 3 Park Ave., 17th Floor,
New York, NY 10016-5997 **TEL: +1 212 419 7555**
BUREAU Palo Alto, Calif.; Tekla S. Perry **+1 650 752 6661**

ADVERTISING INQUIRIES Naylor Association Solutions, Erik Albin +1 352 333 3371, ealbin@naylor.com

REPRINT SALES +1 212 221 9595, ext. 319

IEEE SPECTRUM (ISSN 0018-9235) is published monthly by The Institute of Electrical and Electronics Engineers, Inc. All rights reserved. © 2023 by The Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, New York, NY 10016-5997, U.S.A. Volume No. 60, Issue No. 4. 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, ℅ 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 3401 N. Heartland Dr, Liberty, MO 64068, U.S.A. Printed at Spenta Multimedia Pvt. Ltd., Plot 15, 16 & 21/1, Village Chikhloli, Morivali, MIDC, Ambernath (West), Dist. Thane. IEEE Spectrum circulation is audited by BPA Worldwide. IEEE Spectrum is a member of the Association of Business Information & Media Companies, the Association of Magazine Media, and Association Media & Publishing. IEEE prohibits discrimination, harassment, and bullying. For more information, visit https://www.ieee.org/about/corporate/governance/p9-26.html.

News

California startup Universal Hydrogen is now testing a prototype turboprop plane that's partially powered by two hydrogen fuel cells.

UNIVERSAL HYDROGEN

AVIATION

Hydrogen-Powered Flight Cleared for Takeoff

› The lightest element vies to be jet fuel's future

BY PRACHI PATEL

The 40-seat Dash 8 turboprop plane that took off into the clear, blue, March sky near Moses Lake, Wash., looked ordinary at first glance. Spinning its left propeller, however, was a 1-megawatt electric motor powered by hydrogen fuel cells—the right side ran on a standard Pratt & Whitney turboprop engine—making it the largest aircraft flown on hydrogen to date.

The 15-minute test flight from Hawthorne, Calif.–based Universal Hydrogen came on the heels of a January test flight of a 19-seat Dornier 228 propeller plane by another California startup, ZeroAvia. The CEOs of both companies called these milestones in sustainable aviation "historic." And both companies promise commercial flights of retrofitted turboprop aircraft by 2025. French aviation giant Airbus is going bigger with a planned 2026 demonstration flight of its iconic A380 passenger airplane, which will fly using both hydrogen fuel cells and an engine that burns hydrogen. And Rolls-Royce is making headway on aircraft engines that burn hydrogen.

The aviation industry, responsible for some 2.5 percent of global carbon emissions, has committed to net-zero emissions by 2050. Getting there will require several routes, including sustainable fuels, hybrid-electric engines, and battery-electric aircraft.

Hydrogen is another potential route. Whether used to make electricity in fuel cells or burned in an engine, it combines with oxygen to emit water vapor. If green hydrogen scales up for trucks and ships, it could be a low-cost fuel without the environmental issues of batteries.

Flying on hydrogen brings storage and aircraft certification challenges, but aviation companies are doing the groundwork now for hydrogen flight by 2035. “Hydrogen is headed off to the sky and we’re going to take it there,” says Amanda Simpson, vice president for research and technology at Airbus Americas.

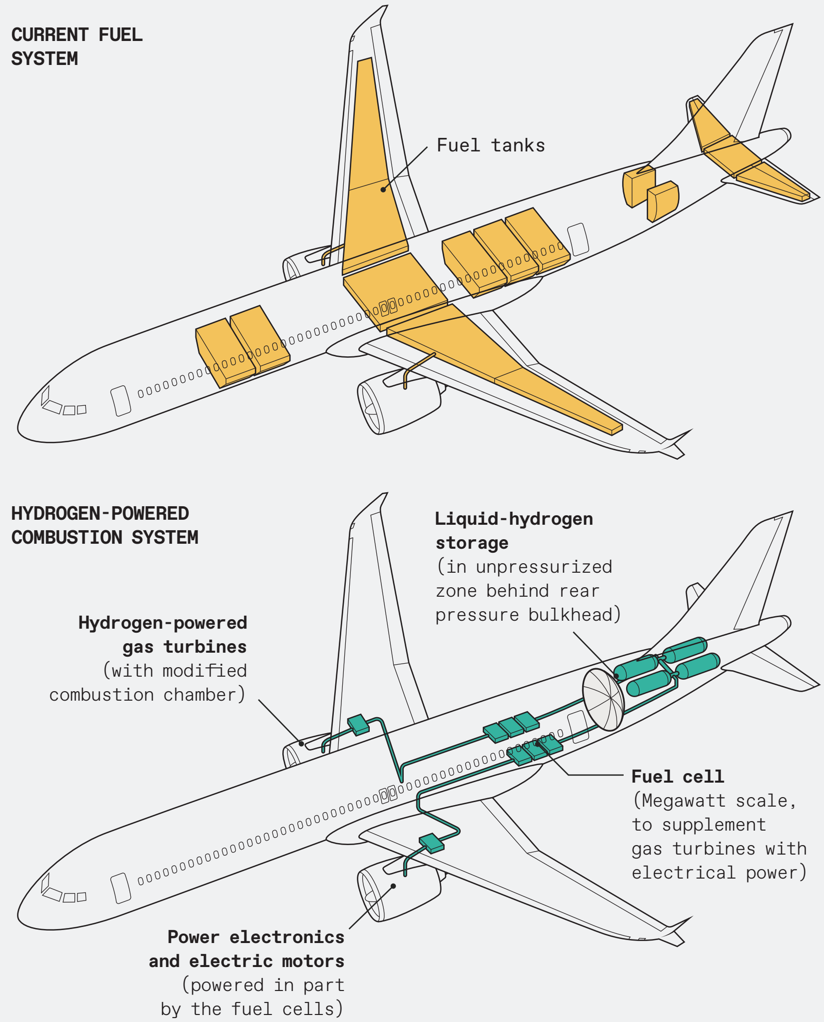
The most plentiful element, hydrogen is also the lightest—key for an industry fighting gravity—packing three times the energy of jet fuel by weight. The problem with hydrogen is its volume. For transport, it has to be stored in heavy tanks either as a compressed high-pressure gas or a cryogenic liquid.

ZeroAvia is using compressed hydrogen gas because it is already approved for road transport. Its test airplane had two hydrogen fuel cells and tanks sitting inside the cabin, but the team is now thinking creatively about a compact system with minimal changes to aircraft design to speed up certification in the United States and Europe. The fuel cells’ added weight could reduce flying range but “that’s not a problem because aircraft are designed to fly much further than they’re used,” says vice president of strategy James McMicking.

The company has backing from investors that include Bill Gates and Jeff Bezos; partnerships with British Airways and United Airlines; and 1,500 preorders for its hydrogen-electric power-train system, half of which are for smaller, 400-kilometer-range 9 to 19 seaters.

By 2027, ZeroAvia plans to convert larger, 70-seater turboprop aircraft with twice the range, used widely in Europe. The company is developing 5-MW electric motors for those, and it plans to switch to more energy-dense liquid hydrogen to save space and weight. The fuel is novel for the aviation industry and could require a longer regulatory approval process, McMicking says.

Next will come a 10-MW power train for aircraft with 100 to 150 seats, “the workhorses of the industry,” he says. Those planes—think Boeing 737—are responsible for 60 percent of aviation



Conventional, fossil-fuel-based airplanes feed fuel from the tanks into the jet engine [above]. A hydrogen-powered system [below] would use liquid or highly compressed hydrogen as fuel that is either burned directly in a combustion engine or converted to electricity by fuel cells that drive an electric motor.

emissions. Making a dent in emissions with hydrogen will require much more efficient fuel cells. ZeroAvia is working on proprietary high-temperature fuel cells for that, he says, with the ability to reuse the large amounts of waste heat generated. “We have designs and a technology road map that take us into jet-engine territory for power,” McMicking says.

Universal Hydrogen, which counts Airbus, American Airlines, and GE Aviation among strategic investors, is placing bets on liquid hydrogen. The startup, “a hydrogen supply and logistics company at our core,” wants to ensure a seamless delivery network for hydrogen aviation as it catches speed, says cofounder and CEO Paul Eremenko. The company sources green hydrogen, turns it into liquid, and puts it in relatively low-tech insulated aluminum tanks that it will deliver via road, rail, or ship. “We want them certified by the Federal Aviation

“Hydrogen is headed off to the sky and we’re going to take it there.”

—AMANDA SIMPSON, AIRBUS AMERICAS

Illustration by James Provost

Administration for 2025, which means they can't be a science project," he says.

Green hydrogen is expected to cost about the same as jet fuel by 2025, Eremenko says. But "there's nobody out there with an incredible hydrogen airplane solution. It's a chicken-and-egg problem."

To crack it, Universal Hydrogen partnered with leading fuel-cell-maker Plug Power to develop a few thousand conversion kits for regional turboprop airplanes. The kits will let operators swap out the engine for a fuel-cell stack, power electronics, and a 2-MW electric motor, all in the existing housing. While the company's competitors use batteries as buffers during takeoff, Eremenko says Universal Hydrogen uses clever algorithms to manage fuel cells so they ramp up and respond quickly. "We are the Nespresso of hydrogen," says Eremenko. "We buy other people's coffee, put it into capsules, and deliver to customers. But we have to build the first coffee machine. We're the only company incubating the chicken and egg at the same time."

Fuel cells have a few advantages over a large central engine. They allow manufacturers to spread smaller propulsion motors over an aircraft, which gives them more design freedom. And because there are no high-temperature moving parts, maintenance costs can be lower. For long-haul aircraft though, the weight and complexity of high-power fuel cells makes hydrogen-*combustion* engines appealing.

Airbus is considering both fuel-cell and combustion propulsion for its ZEROe hydrogen aircraft system. It has partnered with German automotive fuel-cell-maker Elring Klinger, and for direct combustion engines, with CFM International, a joint venture between GE Aviation and Safran. Burning liquid hydrogen in today's engines is still expected to require slight modifications, such as a shorter combustion chamber and better seals.

Airbus is also evaluating hybrid propulsion concepts with a hydrogen-engine-powered turbine and a hydrogen-fuel-cell-powered motor on the same shaft, says Simpson. "Then you can optimize it so you use both propulsion systems for takeoff and climb, and then turn one off for cruising."

The company isn't limiting itself to simple aircraft redesign. Hydrogen tanks could be stored in a cupola on top of the plane, pods under the wings, or a large tank at the back, Simpson says. Without liquid fuel in the wings, as in traditional airplanes, she says, "you can optimize wings for aerodynamics, make them thinner or longer. Or maybe a blended-wing body, which could be very different. This opens up the opportunity to optimize aircraft for efficiency." Certification for such new aircraft could take years, and Airbus isn't expecting commercial flights until 2035.

Conventional aircraft made today will be around in 2050, given their 25- to 30-year life span, says Robin Riedel, an analyst at McKinsey & Co. Sustainable fuels are the only green option for those. He says hydrogen could play a role there through "power-to-liquid technology, where you can mix hydrogen and captured carbon dioxide to make aviation fuel."

Even then, Riedel thinks hydrogen will likely be a small part of aviation's sustainability solution until 2050. "By 2070, hydrogen is going to play a much bigger role," he says. "But we have to get started on hydrogen now." The money that Airbus and Boeing are putting into hydrogen is a small fraction of aerospace spending, he says, but big airlines investing in hydrogen companies or placing power-train orders "shows there is desire."

The aviation industry has to clean up if it is to grow, Airbus's Simpson says. Biofuels are a stepping stone, because they reduce only carbon emissions, not other harmful ones. "If we're going to move towards clean aviation, we have to rethink everything from scratch and that's what ZEROe is doing," she says. "This is an opportunity to make not an evolutionary change but a truly revolutionary one." ■

SEMICONDUCTORS

Big Trouble in Little Interconnects

› At Moore's Law's outer limits, connecting components is complicated

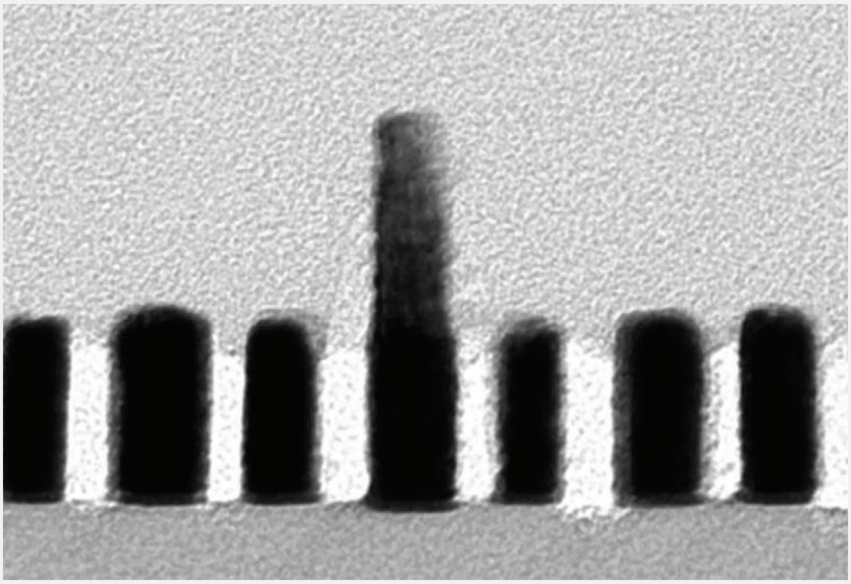
BY SAMUEL K. MOORE

Interconnects—those ultrathin metal wires that link transistors on an IC—are the choke point of an industry that is nearing the outer limits of Moore's Law.

"For some 20 to 25 years now, copper has been the metal of choice for interconnects. However, we're reaching a point where the scaling of copper is slowing down," Chris Penny, lead engineer for advanced interconnect scaling at IBM, said in December at the IEEE International Electron Device Meeting (IEDM). "And there is an opportunity for alternative conductors."

Ruthenium is a leading candidate because, for the smallest cross-section interconnects, it offers better conductivity than copper. But swapping one metal for another isn't easy: These new interconnects will need a different shape and a higher density. And they'll need better insulation, lest signal-sapping capacitance take away all their advantage. Even where the interconnects go is set to change.

The old formulas used to build copper interconnects don't work well with ruthenium. Copper interconnects are built using what's called a damascene process. First, chipmakers use lithography to carve the shape of the intercon-



IBM and Samsung developed a new way to make fine ruthenium interconnects [black] and vertical connections [center pillar]. Importantly, the process allowed for the addition of air gaps [white] between the interconnects.

nect into the dielectric insulation above the transistors. Then they deposit a liner and a barrier material, which prevents copper atoms from drifting out into the rest of the chip to muck things up. Copper, electroplated onto the wafer, then fills the trench. In fact, it overfills it, so the excess must be polished away.

All that extra stuff, the liner and barrier, take up as much as 50 percent of the interconnect volume, Penny said. So the conductive part is narrowing, increasing resistance. But IBM and Samsung researchers have found a way to build tightly spaced, low-resistance ruthenium interconnects that don't need a liner or barrier. The process is called spacer-assisted litho-etch litho-etch, or SALELE, and, as the name implies, it relies on a double helping of extreme-ultraviolet lithography. Instead of filling trenches in the dielectric, it etches the ruthenium interconnects out of a layer of metal and then fills the gaps with dielectric.

The researchers achieve the best resistance using tall, thin, horizontal interconnects. However, that increases capacitance too much. Fortunately, due to the way SALELE builds vertical connections called vias—on top of horizontal interconnects instead of beneath them—the spaces between slender ruthenium lines can easily be filled with air, which is the best insulator available. For these tall, narrow interconnects "the potential benefit of adding an air gap is huge...as much as a 30 percent line capacitance reduction," said Penny.

The SALELE process "provides a road map to 1-nanometer processes and beyond," he said.

By as early as 2024, Intel plans to move power-carrying interconnects from the surface of the silicon to a layer underneath it, so that they approach the transistors from below. Such back-side power delivery has two main advantages: It allows electricity to flow through wider, less resistive interconnects, reducing power loss. And it frees up room above the transistors for signal-carrying interconnects, meaning logic cells can be smaller. (See "Power From Below," *IEEE Spectrum*, September 2021.)

At IEDM 2022, Imec researchers announced ways to move the end points of the power-delivery network closer to transistors without messing up the electronic properties of those transistors. But they also reported a somewhat troubling problem: Back-side power could lead to a buildup of heat when used in a 3D stack of chips.

First the good news: Imec researchers found that the transistors worked just fine even when the power-delivery rail was built right alongside the transistor channel region (though still tens of nanometers below it). And that could mean even smaller logic cells.

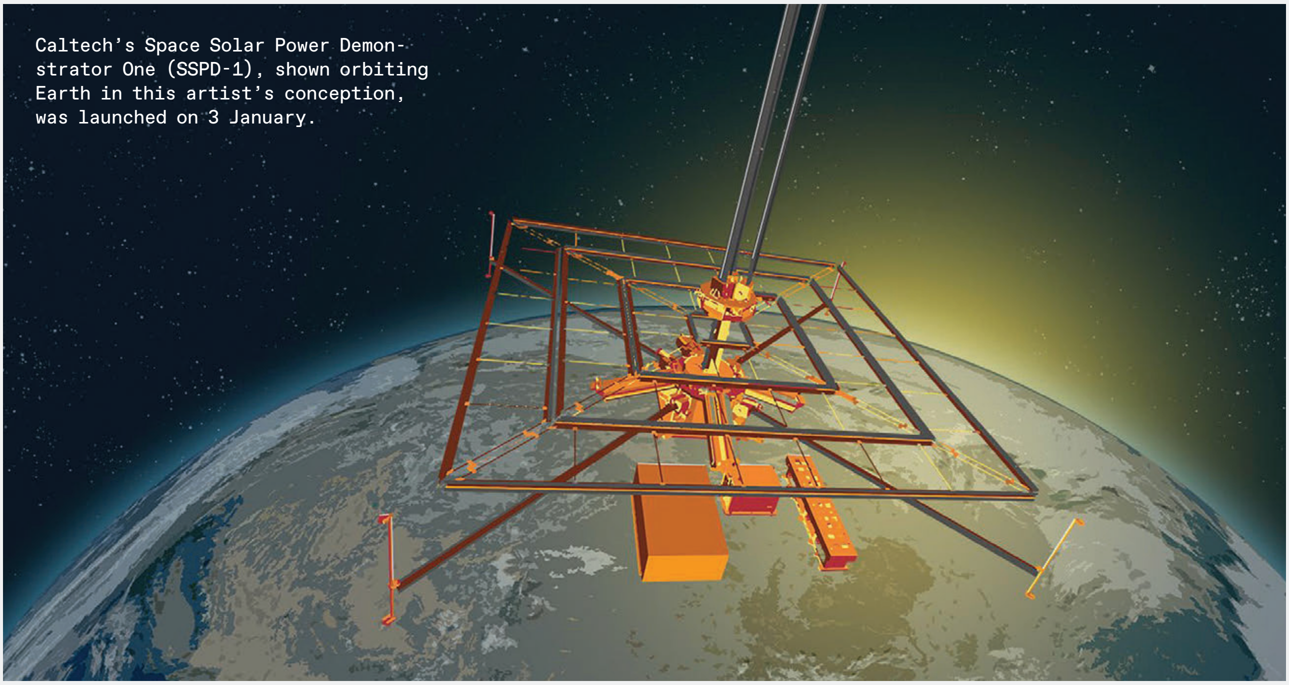
Now the bad news: In separate research, Imec engineers simulated several versions of the same future CPU and found some troubling hot spots. Some of the simulations had today's front-side power delivery, where interconnects for both data and power are built in layers above the silicon. Some had back-side power delivery networks. And one was a 3D stack of two CPUs bonded face to face, the bottom having back-side power and the top having front-side.

Back-side power's advantages were confirmed by the simulations of the 2D CPUs. Compared with front-side delivery, it cut the loss from power delivery in half. And transient voltage drops were less pronounced. What's more, the CPU area was 8 percent smaller. However, the temperature difference between the hottest part of the back-side chip and its edge was about 45 percent greater than the difference between the hottest part of a front-side chip and its edge. The likely cause is that back-side power requires thinning the chip down to the point where it needs to be bonded to a separate piece of silicon just to remain stable. That bond acts as an insulator, trapping heat inside the chip.

The real problems arose with the 3D IC. The top CPU has to get its power from the bottom CPU, but the long journey to the top had consequences. While the bottom CPU still had more stable voltage levels than a front-side chip, the top CPU performed much worse in that respect. And the 3D IC's power network ate up more than twice the power that a single front-side chip's network would consume. Worse still, heat could not escape the 3D stack very well, with the hottest part of the bottom die almost 2.5 times as hot as a single front-side CPU. The top CPU was cooler, but not by much.

The 3D IC simulation is admittedly somewhat unrealistic, Rongmei Chen, back-side power network project leader and senior research scientist at Imec told engineers at IEDM. Stacking two otherwise identical CPUs atop each other is an unlikely scenario. (It's much more common to stack memory with a CPU.) "It's not a very fair comparison," he said. But it does point out some potential issues. ■

IBM



Caltech's Space Solar Power Demonstrator One (SSPD-1), shown orbiting Earth in this artist's conception, was launched on 3 January.

CALTECH

ENERGY

Trial Run for Orbiting Solar Array › Caltech satellite tests out solar power beamed from space

BY NED POTTER

For about as long as engineers have talked about beaming solar power to Earth from space, they've cautioned that it was unlikely to happen anytime soon. Photovoltaic cells were inefficient, so orbiting arrays would need to be the size of cities. The plans got no closer to space than the upper shelves of libraries.

That's beginning to change. On 3 January, a SpaceX Falcon 9 rocket launched a small experimental satellite into a polar orbit about 525 kilometers high, where the spacecraft is always in sunlight. It's called the Space Solar Power Demonstrator One (SSPD-1), and it was designed and built at the California Institute of Technology and funded by donations from California real estate developer Donald Bren.

"To the best of our knowledge, this would be the first demonstration of actual power transfer in space, of wireless power transfer," says Ali Hajimiri, a professor of electrical engineering at Caltech and a codirector of the program behind SSPD-1, the Space Solar Power Project.

If all goes well, SSPD-1 will spend at least five months testing prototype components of possible future solar stations in space. The project managers hope to test a lightweight frame, called DOLCE (Deployable on-Orbit ultra-Light Composite Experiment), on which future solar arrays could be mounted. Another assembly on the spacecraft contains samples of 32 different types of photovoltaic cells, intended to see which would be most efficient and robust. A third part contains a microwave transmitter, set up to prove that energy from the solar cells can be sent to a receiver. For this first experiment, the receivers are on board the spacecraft. But if it works, an obvious future step would be to send electricity via microwave to receivers on the ground.

The Caltech experiment may be happening in the near vacuum of space, but here on Earth there is a rush of activity. Airbus, the European aerospace company, has been testing its own technology on the ground, and government agencies in China, Japan, South Korea, and the United States have all mounted small projects. "Recent technology and conceptual advances have made the concept both viable and economically competitive," said Frazer-Nash, a British engineering consultancy, in a 2021 report to the U.K. government.

For one thing, the cost of launching hardware into orbit keeps dropping, led by SpaceX and other, smaller companies such as Rocket Lab. SpaceX has a simplified calculator on its website, showing that if you want to launch a 50-kilogram satellite into sun-synchronous orbit, the company will do it for US $333,000.

Space launches using the European Ariane 5 rocket beginning in 2003 cost roughly 40 percent more.

Meanwhile, photovoltaic technology has improved. Electronic components keep getting better and cheaper. And there is political pressure as well: Governments and major companies have made commitments to decarbonize in the battle against global climate change, committing to renewable energy sources to replace fossil fuels.

Most solar power, at least for the foreseeable future, will be Earth-based, which will be cheaper and easier to maintain than anything anyone can launch into space. Proponents of space-based solar power say that for now, they see it as best used for specialty needs, say to supply remote outposts, places recovering from disasters, or even other space vehicles.

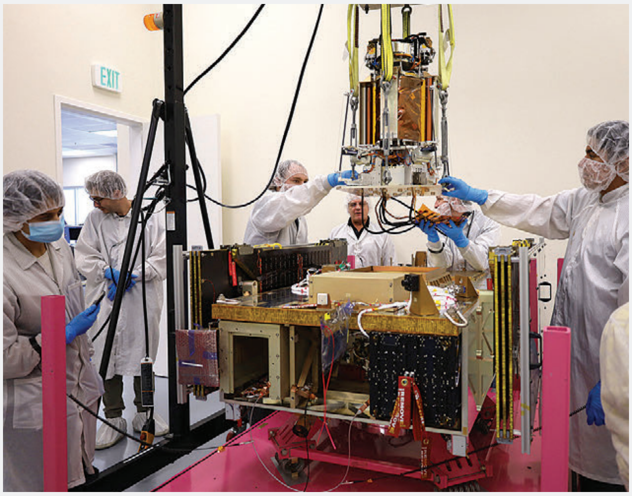
But don't underestimate the advantages of space, Hajimiri says. The unfiltered sunlight up there is considerably stronger than what we see down on the ground, and it's uninterrupted by darkness or bad weather—if you can build an orbiting array that is light enough to be practical.

Past designs were constrained by the technology of their times. Engineers envisaged impossibly large truss structures to hold solar panels and wiring to route power to a central transmitter. The Caltech team would dispense with all that. An array would consist of thousands of tiles, each as small as 100 square centimeters, each with its own solar cells, transmitter, and avionics. They might be loosely connected, or they might even fly in formation.

"The analogy I like to use is that it's like an army of ants instead of an elephant," says Hajimiri. Transmission to receivers on the ground could be by phased array—a cascade of microwave signals, each from a different tile, emitted in phases that have been synchronized so that they form a beam that can be aimed without having recourse to any moving parts.

"That's the kind of thing we're talking about," said Harry Atwater, a coleader of the Caltech project, as SSPD-1 was being planned. "Really gossamer-like, ultralight, the limits of mass-density deployable systems."

If it works out, maybe fleets of orbiting solar power stations will add to the world's energy mix in the coming decades. As a 2022 report from Frazer-Nash concluded, this is "a potential game changer." ■



Caltech engineers practice the procedure for loading a component of the SSPD-1 onto the launch vehicle using scale models.

JOURNAL WATCH

Home Water Heaters for Energy Storage

In the modern, renewably powered home, a water heater could also function as a battery.

The concept is simple but effective: When energy is plentiful, it can be stored as heat in the water, meeting demand for hot water later on. A study published in January in *IEEE Open Access Journal of Power and Energy* shows that water heaters are an overall more cost-efficient way to store energy.

Mahan Mansouri, an energy market analyst at Pacific Gas and Electric, conducted the study as part of his Ph.D. dissertation at Ohio State University.

"An individual water heater represents a very small electricity demand on the scale of a bulk electricity system," Mansouri says. "However, aggregating thousands or millions of water heaters could be very low cost compared to deploying batteries." There are already more than 58 million households in the United States with electric water heaters that could be used for this purpose, he adds.

But how does the cost compare with storing energy in 10-megawatt lithium-ion batteries?

For their analysis, Mansouri and his coauthor used real-world data from California, since California rooftop solar panels often generate more energy than can be used. "Thus, having the ability to store that energy midday and use it later during the evening when solar output falls would be of great value," Mansouri says.

The results show that, because batteries can store energy for longer periods of time, they are more profitable in terms of selling energy back into the grid, yielding an annual operating profit almost twice that of water heaters. But batteries cost more to install, and therefore water heaters have a better overall cost-to-profit ratio. **—Michelle Hampson**

CALTECH

Uravu Labs cofounders Govinda Balaji [left] and Swapnil Shrivastav pose next to their company's device for harvesting freshwater from thin air.

ENVIRONMENT

Air, the New Freshwater Source › An Indian startup's technology could ease pressing shortages

BY EDD GENT

Technology that can pull water out of thin air could help solve the world's growing water-scarcity problem, but most solutions are difficult to scale. Indian startup Uravu Labs now offers a low-cost modular unit for atmospheric water harvesting. It's renewably powered, harnesses a vast and nearly inexhaustible source, and it produces no wastewater.

The company says its biggest unit to date, at its headquarters in the South Indian city of Bengaluru, will harvest up to 1,000 liters of water a day. By the end of the year, the company hopes to scale that up to 10,000 L a day, says cofounder Swapnil Shrivastav.

Most companies working in this area rely on technology like that found in air-conditioning units: A coiled tube full of refrigerant is used to cool air until its moisture condenses on the surface as liquid water. That approach requires huge amounts of electricity. "Our goal from day one was to not only be scalable and renewable but also be the most affordable," Shrivastav says.

The key to Uravu's technique is the use of desiccants, which absorb moisture from the air. The material then releases the water again when heated. The researchers started work on the idea in 2017 and made it to the finals of the Water Abundance XPrize, a two-year competition organized by XPrize to catalyze progress in atmospheric water-harvesting technology. Since then, Uravu has raised US $2.5 million.

Initially, the company worked on a self-contained unit, about 4 meters across, that used a solar thermal panel to warm silica gel, a solid desiccant. That version produced about 10 L of water a day. But because each unit required dedicated fans, valves, and pumps, it cost too much to scale up.

So Uravu has now switched to a new design that decouples the harvesting of water vapor and the production of liquid water. In this design, a liquid desiccant—calcium chloride solution—is pumped between dedicated absorber and desorber units. In the absorbers, the desiccant is sprinkled through a mesh to increase the surface area exposed to air, which is drawn over it by a fan. The desiccant then flows to a separate desorber unit where, to release the moisture, it is heated to between 60 and 70 °C by a coiled pipe filled with hot water. The ultrahumid air then passes to a low-power, air-cooled condenser that turns it into liquid water.

Decoupling the absorber and desorber leads to significant scaling efficiencies, says Shrivastav, since each desorber can serve several of the less complicated absorbers. Each absorber unit is capable of collecting as much as 200 L per day. To create Uravu's 1,000-L-per-day machine, the company will couple six absorbers with a single desorber unit. The new design can now use a variety of heat sources, including solar thermal panels, biomass burners, or industrial waste heat. If working with solar thermal power, the design also incorporates a hot water tank that allows it to run even when the sun goes down. And, if the air is notably dry, the operator can simply add more desiccant to get the desired output—an option that conventional systems don't offer.

However, Uravu's system has a big footprint. Competing products that produce 6,000 L per day can fit into a shipping container, but in the near term, Uravu is unlikely to get more than 2,000 L out of a system that size, Shrivastav says.

Still, Uravu's condenser requires roughly 40 percent less power than

EDD GENT

conventional systems and thus produces water for 40 percent less money—about 6 U.S. cents a liter. The company expects the cost will soon come down to 3 cents per liter.

But while it may be cheaper than other air-to-water approaches, Uravu still can't compete with water from conventional sources, which typically costs less than 1 cent per liter in India. The math is more favorable in other geographies with more expensive water, such as the Middle East and Australia, says Shrivastav. The only remaining problem is the necessarily large upfront investment, a feature of all forms of water harvesting, says Chiara Neto, professor of physical chemistry at the University of Sydney. That may put the technology out of reach for many poorer communities most in danger of water scarcity.

Closer to home, Uravu has had to explore some novel business models that involve people paying a premium to burnish their environmental credentials.

The company is currently providing renewable drinking water in reusable glass bottles to high-end restaurants and hotel chains in Bengaluru. Shrivastav says the company's aim is to compete with premium mineral waters that sell at a higher margin than standard drinking water. Uravu has also installed one of its units at India's third-largest distiller, Radico Khaitan, which plans to use its water in high-end spirits as a marketing ploy.

COMPUTING

Deepfakes Are Now an Industry

In December, as many people prepared to take off for a winter vacation, the creators of the animated TV show "South Park" announced US $20 million dollars of new funding for the duo's deepfake studio, Deep Voodoo. The studio, they say, offers "unrivaled face-swapping visual effects to artists, producers, and creators around the world." Dozens of startups now provide services based on deepfake technology. The largest of these, London-based Synthesia, lets users choose from over 100 AI avatars [right] to create videos without a live presenter.

The future of deepfakes is still being written, but it seems to include a wide range of video-manipulation tools that let editors change every aspect of a video. And isn't that, at some point, just video editing?

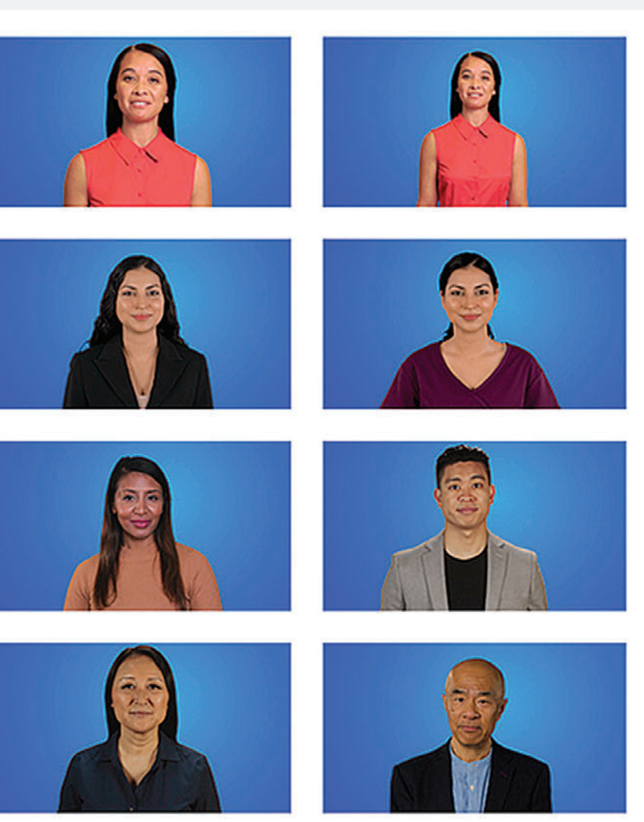
—MATTHEW S. SMITH

Shrivastav is also hopeful that NGOs or government organizations may help subsidize some of the cost of Uravu's units, based on their knock-on benefits. Because they don't rely on natural water sources, the units can be set up anywhere, so people don't have to walk far to collect water. And in India, which has a worsening groundwater-depletion crisis, they could help solve the growing problem of delivering potable water in sufficient quantities.

What water-harvesting systems really need is a better desiccant, says Swee Ching Tan, an assistant professor of materials science and engineering at the National University of Singapore. "We need material science to make more effective, efficient materials" that can hold more water and release it at lower temperatures, he says.

Shrivastav says Uravu should be able to work with new desiccants as they become available. But he points out that new materials typically take 8 to 12 years to reach commercialization. So, waiting for better desiccants isn't an option. And ultimately, even if Uravu's costs limit where the technology can be applied, it could still make a significant difference.

"We want to reach around 1 million L per day in the next five years," he says. "That translates to saving more than 2.5 million L of groundwater and more than 15 to 20 tonnes of CO_2." ■



SYNTHESIA

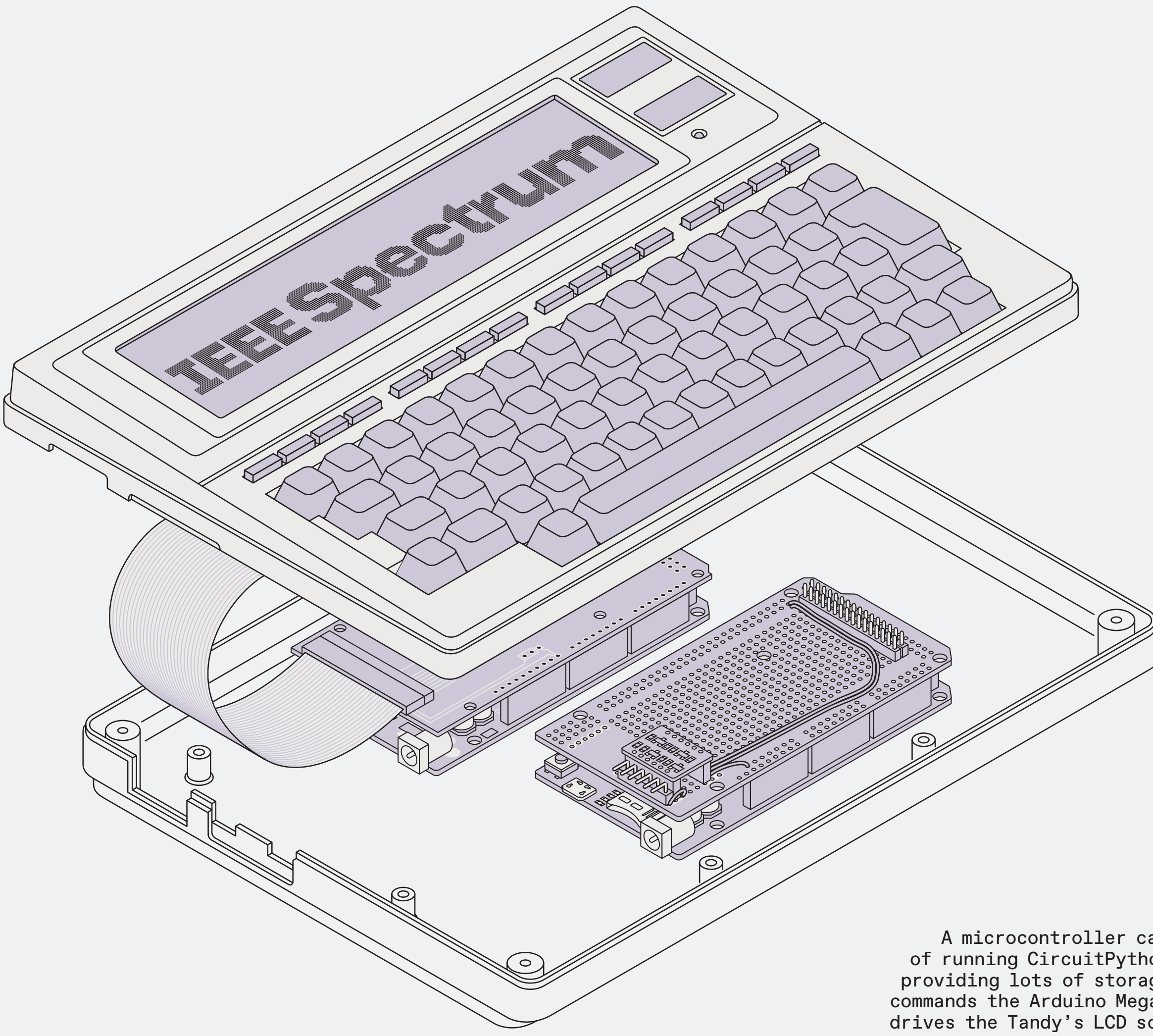
JUICE BOX

By Willie D. Jones

For many years, environmentalists have looked forward to the coming of net-zero-energy buildings. Much effort has been devoted to making lighting, heating, and cooling more efficient so buildings consumed less energy. But the net-zero target would never have been reachable without innovations in renewable-energy generation that let structures generate power on-site. Now, residential and commercial buildings can be outfitted with roofing tiles that double as solar panels, or with rooftop boxes like this low-profile unit, which transforms gusts of wind into electric current. This WindBox turbine, installed on the roof of a building in Rouen, France, is 1.6 meters tall, and has a 4-square-meter footprint (leaving plenty of space for solar panels or tiles). The unit, which weighs 130 kilograms, can generate up to 2,500 kilowatt-hours of electricity per year (enough to meet roughly one-quarter of the energy needs of a typical U.S. household).

PHOTOGRAPH BY LOU BENOIST/ AFP/GETTY IMAGES

Hands On



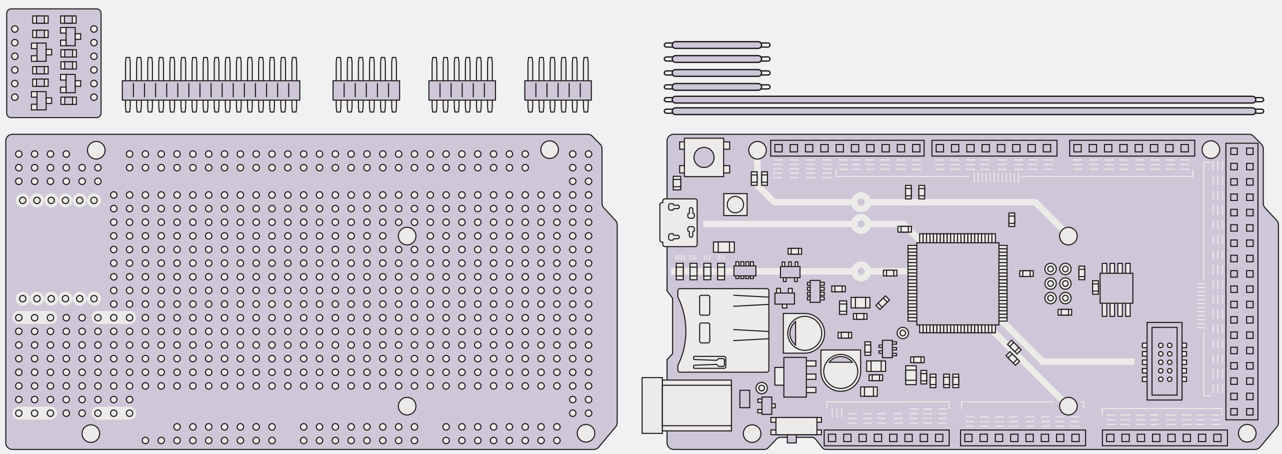
A microcontroller capable of running CircuitPython and providing lots of storage now commands the Arduino Mega that drives the Tandy's LCD screen.

I^2C Strikes Back › Surprises can lurk in even a simple comms protocol

BY STEPHEN CASS

Last year for Hands On, I gutted a defunct TRS-80 Model 100. The goal was to upgrade its 24 kilobytes of RAM and 2.4-megahertz, 8-bit CPU but keep the notebook computer's lovely keyboard and LCD screen. That article was almost entirely about figuring out how to drive its squirrely 1980s-era LCD screen. I left the rest, as they say, as an exercise for the reader. After all, sending a stream of data from a new CPU to the Arduino

Illustrations by James Provost



The Grand Central controller is now the new brains of the Tandy. Although the controller has the same form factor as the Arduino Mega, it has vastly more compute power. A custom-built shield [left] holds a supporting voltage-level shifter [top left] that appropriately converts the 3.3- and 5-volt logic levels used by the controllers.

Mega controlling the screen would be a trivial exercise, right?

Hahahaha.

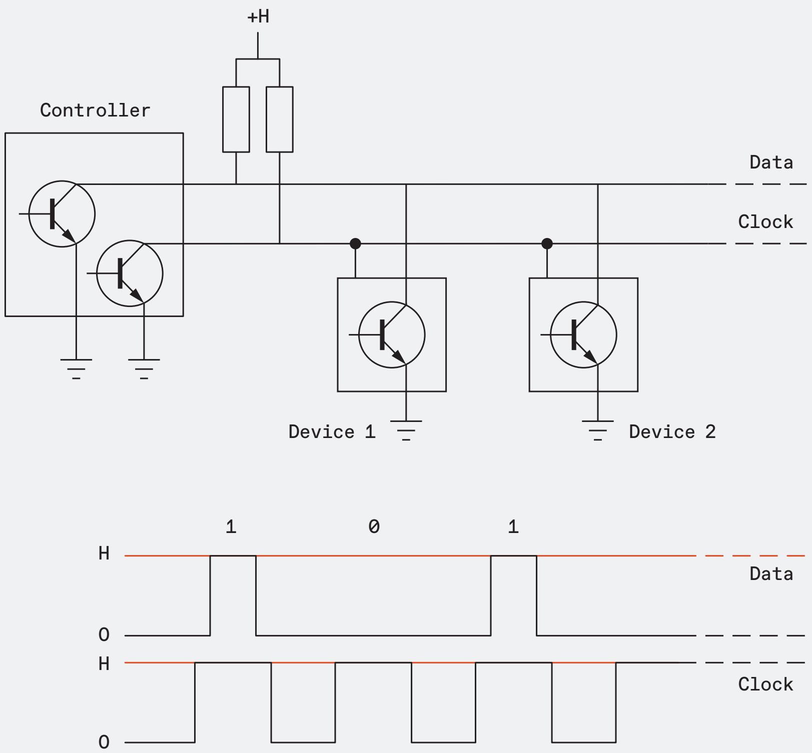
No, folks, no it was not. *IEEE Spectrum's* Hands On articles provide necessarily linear versions of how projects come together. It can give the impression we're terribly clever, which has about the same relationship to reality as an influencer's curated social-media feed. So every now and then I like to present a tale steeped in failure, just as a reminder that this is what engineering's like sometimes.

To send screen data to the Mega, I had a choice between several methods that are supported by CircuitPython's display driver libraries. I wanted to use a CircuitPython-powered microcontroller as the Model 100's new brain because there's a lot of existing software I could port over. In particular, CircuitPython's display libraries would greatly simplify creating graphics and text and would automatically update the display. My choices were between a parallel interface and two serial interfaces: SPI and I^{2}C.

The parallel interface would require at least 12 wires. SPI was better, being a four-wire interface. But I^{2}C was best of all, requiring only two wires! Additionally, there are many breakout boards that support I^{2}C, including storage and sensors of all types. One I^{2}C bus can, in theory, support over a hundred I^{2}C peripherals. I^{2}C is much slower than SPI, but the Model 100's delightfully chunky 240-by-64-pixel display is slower still. And I'd used I^{2}C-based peripherals many times before in previous projects. I^{2}C was the obvious choice. But there's a big difference between using a peripheral created by a vendor and building one yourself.

On the circuit level, I^{2}C is built around an "open drain" principle. When the bus is idle, or when a 1 is being transmitted, pull-up resistors hold the lines at the voltage level, indicating a logical high. Connecting a line to ground pulls it low. One line transmits pulses from the central controller as a clock signal. The other line handles data, with one bit transmitted per clock cycle. Devices recognize when traffic on the bus is intended for them because each has a unique 7-bit address. This address is prepended to any block of data bytes being sent. In theory, any clock speed and/or logic-level voltage could be used, as long as both the controller and peripheral accept them.

And there was my first and, I thought, only problem: The microcontrollers that ran CircuitPython and were computationally hefty enough for my needs ran on 3.3 volts, while the Arduino Mega uses the 5 V required to drive the LCD. An easy solve though—I'd just use a US $4 off-the-shelf logic-level shifter, albeit a special type that's compatible with I^{2}C's open-drain setup.



An I^{2}C is a relatively low-speed bus that provides bidirectional communications between a controller and (in theory) over a hundred peripherals. A data and clock line are kept at a high voltage by pull-up resistors. The frequency of a clock line is established by the controller, while both the control and peripheral devices can affect the data line by connecting it to ground. A peripheral will take control of the data line only after it has been commanded to do so by the controller to avoid communication collisions.

Using a $40 Adafruit Grand Central board as my central controller, I connected it to the Mega via the level shifter, and put some test code on both microcontrollers. The most basic I^{2}C transaction possible is for the controller to send a peripheral's address over the bus and get an acknowledgment back.

No response. After checking my code and wiring, I hooked up a logic analyzer to the bus. Out popped a lovely pulse train, which the analyzer software decoded as a stream of correctly formed addresses being sent out by the Grand Central controller as it scanned for peripherals, but with no acknowledgment from the Mega.

I'll skip over the next few hours of diagnostic failure, involving much gnashing of teeth.

I'll skip over the next few hours of diagnostic failure, involving much gnashing of teeth and a dead end involving a quirk in how the SAMD chip at the heart of the Grand Central controller (and many others) has a hardware-accelerated I^{2}C interface that reportedly can't go slower than a clock speed of 100 kilohertz. Eventually I hooked up the logic analyzer again, and scrolling up and down through the decoded pulses I finally noticed that the bus scan started not at address 0 or 1, but at 16. Now, when I'd picked an address for the Mega in my test code, I'd seen many descriptions of I^{2}C in tutorials that said the possible range of addresses ran from 0 to 127. When I'd looked at what seemed like a pretty comprehensive description by Texas Instruments of how the I^{2}C bus worked down to the electrical level, addresses were simply described as being 7 bit—that is, 0 to 127. So I'd picked 4, more or less at random.

But with the results of my logic scan in hand, I discovered that, *oh, by the way,* addresses 0 to 7 are actually unusable because they are reserved for various bus-management functions. So I went back to my original hardware setup, plugged in a nice two-digit address, and bingo! Everything worked just fine.

True, this headache was caused by my own lack of understanding of how I^{2}C works. The caveat that reserved addresses exist *can* be found in some tutorials, as well as more detailed documentation from folks like Texas Instruments. But in my defense, even in the best tutorials it's usually pretty buried and easy to miss. (The vast majority of I^{2}C instruction concerns the vastly more common situation where a grown-up has built the peripheral and hardwired it with a sensible address.) And even then, nothing would have told me that CircuitPython's heartbeat scan would start at 16.

Oh well, time to press on with the upgrade. The rest should be pretty easy, though! ■

BY COLIN ASHRUF

Careers: 10 Tips for Product Development

Getting to market requires more than clever engineering

NITAT TERMMEE/GETTY IMAGES

For more than two decades I have been a consultant for the high-tech and medical-device industries. I'm typically brought in to advise companies on the management of troubled, underperforming projects and engineering operations.

Here are 10 challenges I frequently come across in product development, along with ways to mitigate or avoid them.

1. Nail down interfaces

A clear definition of the scope of your project is essential to establish exactly where these boundaries are. Don't assume the scope is clear to all; it's not. Clarify, write it down, and share it with the teams and stakeholders—and do so before development work begins.

Describe in detail all system and subsystem interfaces, whether physical (hardware) or nonphysical (such as software or indirect couplings between a system and its local environment). Assign interface managers to keep track of demarcations and manage all the related agreements and issues. This is especially relevant when you and your team need to cooperate with development partners.

It takes courage to terminate a project—that's why many dead-on-arrival projects drag on.

2. Manage relationships with temporary workers

Project teams often include temporary workers, whether on contract or from partner organizations. A common approach is to start off the relationship with these workers by specifying their subprojects and trying to set unambiguous deliverables and deadlines. But what these often vague agreements actually do is trigger endless arguing about contract variations, project scope, and cost escalations, which leads to mistrust and delays.

A more productive approach is to integrate temporary workers into the project and treat them as part of your organization. Integrated teams are happier, more motivated, and more productive. The end result: faster product development.

3. Insist on realistic quality management

Ensuring quality in products is painstaking work that involves following carefully thought-out processes and taking care of the associated documentation—the stuff that few people like to do.

Do you have business processes in place for developing and launching new products, including an active verification mechanism?

"We have a quality-management system" is the frequent answer. But do you and your team understand and apply it? Does it satisfy present-day business needs, or is it full of boilerplate stuff that you do just because that's the way you always did it?

A quality-management system whose description spans 800 pages across 100 documents is not workable. Keep it short, 50 pages max. And don't let the rules become a straitjacket. If you need to deviate from the quality-management system for good reasons, do so—but document the changes and revise the rules to match current needs.

4. Don't confuse planning with scheduling

Many people conflate planning and scheduling, but they are not the same. Planning comes first: What do you need to do to finish the project? The plan is a summary of all the steps. Make it explicit, and list all the work-breakdown elements, work packages, backlog user stories, and other steps. Don't worry if every detail isn't clear; refine as you go. This forces you to think up front about the challenges and plan for their resolutions, including skills, staffing, resources, and budgets.

Once you have a plan, don't just add dates to create a schedule and drop it on your staff. Those who carry out the work need to give input on their lead times. If you set dates for them, you'll miss the mark, and the team won't feel committed to the schedule. Refrain from inserting buffers everywhere; just make the schedule as realistic as possible. The schedule is a best-guess estimate, and everyone will need to treat it as such. Missed deadlines should be analyzed and their reasons discussed in retrospective sessions.

5. Train the whole team

Project management is not the sole responsibility of the project manager. Projects are managed by the entire team. If you add a great project manager to a team that cares nothing for structure, you'll end up with chaos. That's why I'm a firm believer in training the entire team, down to the most junior member, on project-control techniques.

6. Separate research from development

Don't incorporate research into your product-development projects. Product development should be about combining state-of-the-art—though proven—technology in new ways. Don't try to innovate by incorporating unproven technology. This doesn't mean a company shouldn't try to break new ground or come up with wild ideas. Just acknowledge that unproven technology often takes years to mature.

It's okay to have a team work on research alongside development. But first, separate the research from the regular development work and mark the research items as such so the associated risks and challenges remain transparent to all. If the research doesn't pan out, you'll save yourself the resources you would otherwise have wasted on dead-end development.

7. Demos are not enough

Demonstrations are a great way of getting the team and stakeholders updated and involved. But their use is no excuse for skipping documentation or failing to understand your product and the underlying science.

8. What a steering committee should—and shouldn't—do

A steering committee exists to help with impediments that developers and project managers cannot resolve. This includes navigating roadblocks within the organization and mitigating high-level issues, including capacity or personnel availability, requests for more money, and conflicts with strategic partners or suppliers.

What a steering committee should not do is "help" the project team by doing its work. If a committee member has a strong opinion about, say, the way the engineers broke down their electronics design or set up their code base, then they should join the project team. Otherwise, they are holding up the project, creating frustration, and possibly even sinking the entire endeavor.

Ensuring quality in products is painstaking work that involves following carefully thought-out processes and taking care of the associated documentation.

9. Don't get hung up analyzing which stage a project is in

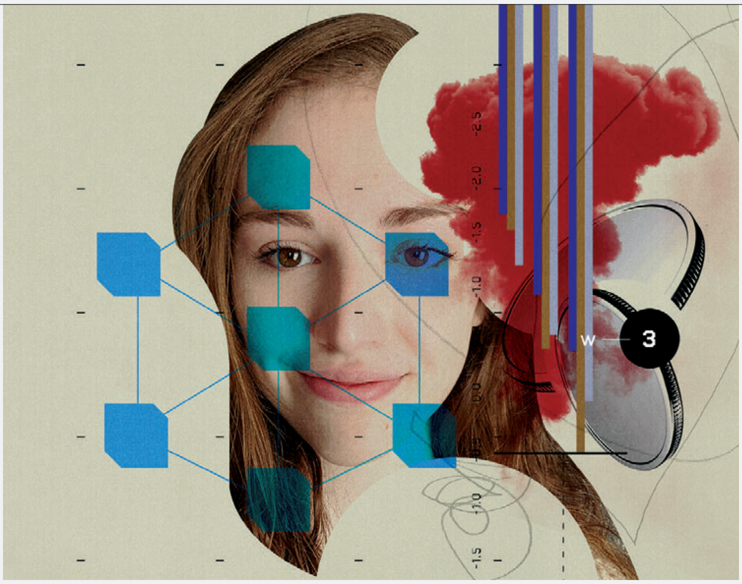
Stages are used to avoid wasted effort and resources on products that won't make it to market, creating checkpoints where would-be flops can be terminated early. That's why all variations of staged development go from high risk/low cost to low risk/high cost.

Initially, you try to eliminate the risks of not making it to market. In the final stages, you pour money and resources into building up the production lines and setting up the field sales and service organizations that will support the market. Whether the project is in phase two or phase three of 10 is really not a big deal. Avoid senseless arguments: Fix any missing deliverables associated with a phase at a later stage, document the deviations from the process, and move on.

Project staging enables making go/no-go decisions early. Giving the go is always easier. It takes courage to terminate a project—that's why many dead-on-arrival projects drag on. Be bold if the signs are mostly red, and be bold as early as possible.

10. Product development is for marathoners, not sprinters

Product development can be painstaking drudgery at times. Your team has to work through failures and setbacks, and spend long hours on the factory floor getting the design successfully into production. Make sure you have the right people for the job. If someone complains that the work is beneath their training or intellect, take them off the project. Product development is for diehards—those who don't complain and keep at it no matter what and understand that even the most exciting jobs in the world cannot be interesting all the time. The mentality, the will to do it, and a healthy dose of curiosity are more important than an impressive résumé or fancy title. ■



5 Questions for Molly White

The "Web3 Is Going Just Great" creator on why it isn't

In mid-2021, the term "Web3" suddenly exploded into the public consciousness. As people scrambled to figure out what it was—cryptocurrencies? blockchain? nonfungible tokens?—venture capital firms were pouring money into new startups, totaling over US $30 billion before the year was out.

Meanwhile, Molly White, a software engineer, started reading up on the tech in case that was the direction her career would be heading in. But she found herself taking a different direction: She launched the website Web3 Is Going Just Great, with the aim of tracking the scams and fraud in the cryptocurrency world. So far, she's tallied $11.8 billion in money lost on the website's Grift Counter. White answered five rapid-fire questions on the Web3 phenomenon and why she's still not impressed.

How did you end up running a site like Web3 Is Going Just Great?
Molly White: When I started researching the topic, I was just seeing a lack of reporting on some of the downsides—you know, the hacks, the scams, the fraud. And so I decided I could do my part to try and fill that void to some extent, because I feel like it's important that people get the full picture.

Molly White is a software engineer and the creator of the Web3 Is Going Just Great website. She is also a fellow at the Harvard Library Innovation Lab.

A lot of the projects you're tracking involve cryptocurrency and blockchain technologies. Is that what "Web3" means? Are all of these terms synonymous?
White: It's primarily a marketing term. And I think the industry benefits from how nebulous it is because it can mean whatever is most useful at that time. But broadly speaking, Web3 refers to blockchains underpinning everything you do online.

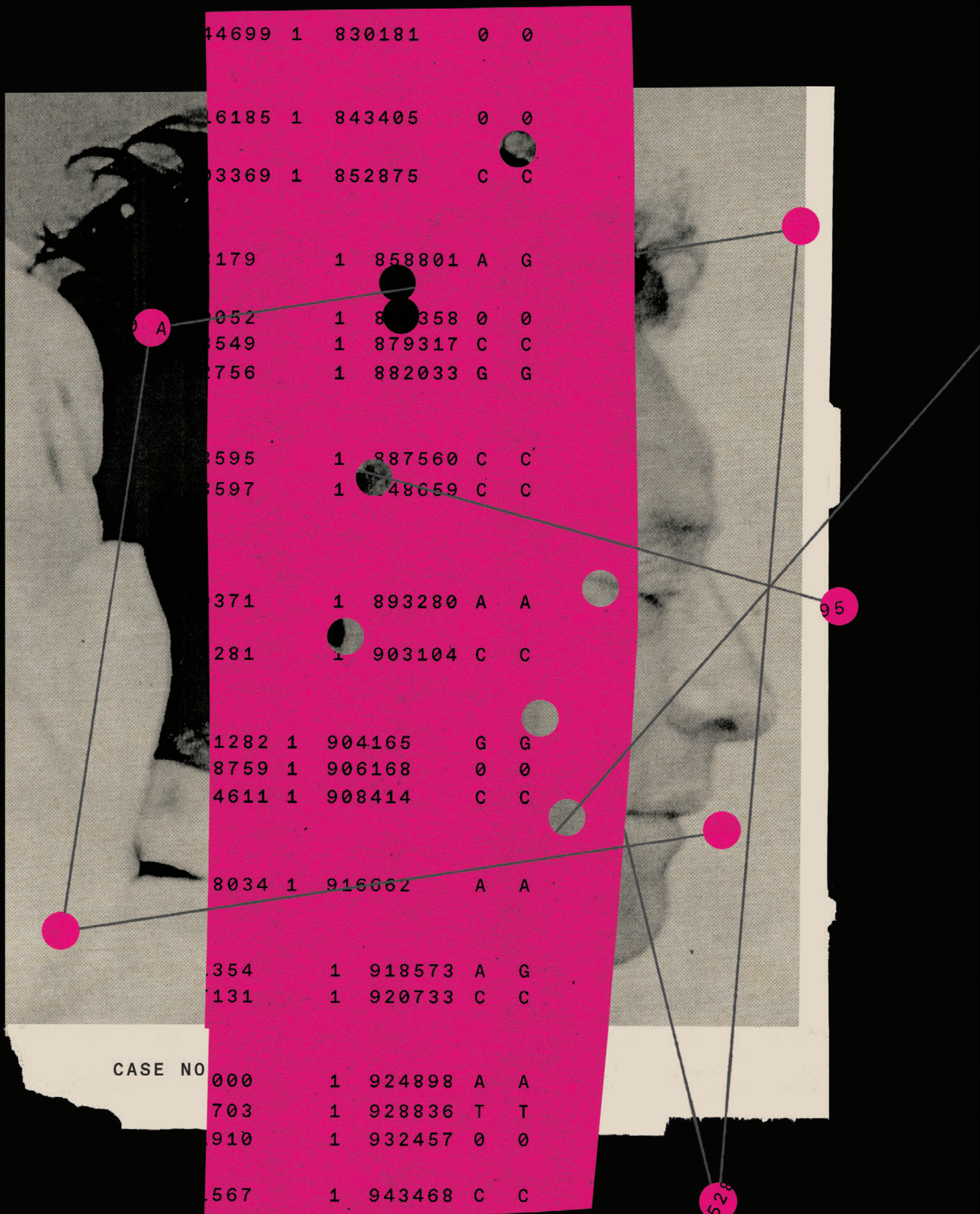
The crypto industry seemed like it might collapse when the cryptocurrency exchange FTX went bankrupt in November 2022, but you're still updating the site with new projects. Is the industry still just trucking along, or has it changed after that event?
White: I think that FTX and the related collapses have been a really big hit to the crypto "brand," but I think that the crypto industry is constantly working on finding the next big thing that they can sell retail investors on. And so that is very much underway at this point.

You can sort of see what's happening as people start distancing themselves from FTX and saying that the FTX collapse wasn't a flaw of crypto—it was a flaw of centralization or fraudulent actors. So I get the sense that people are going to be moving toward selling people on more decentralized finance products. That's my guess of what the next big thing is going to be. It's either that or crypto meets AI. We'll see.

Have you ever come across a project that made you think, "Oh, maybe there's a worthwhile reason for adding a blockchain to this"? Or are you still waiting for that project?
White: I'm mostly still waiting. Every once in a while there's something where I can understand what they're going for, but I don't understand why they've picked a blockchain over a more efficient or less expensive solution. And sometimes there's individual cases where people have benefited from crypto, but I don't necessarily see that as scalable, or a strong argument for the technology itself.

Do you think you're more skeptical about crypto and Web3 than when you started Web3 Is Going Just Great?
White: Well, I still have an open mind. I still tell people that I'm open to there being some killer use case that I just haven't thought about. But seeing the constant fraud—and how motivated people are by the economic forces in crypto to take advantage of people—has really made me very skeptical and cynical about the industry. ■

Photo-illustration by Stuart Bradford



44699 1 830181 0 0
6185 1 843405 0 0
3369 1 852875 C C
179 1 858801 A G
A
052 1
358 0 0
549 1 879317 C C
756 1 882033 G G
595 1 887560 C C
597 1 48659 C C
95
371 1 893280 A A
281 903104 C C
1282 1 904165 G G
8759 1 906168 0 0
4611 1 908414 C C
8034 1 916662 A A
354 1 918573 A G
131 1 920733 C C
CASE NO
000 1 924898 A A
703 1 928836 T T
910 1 932457 0 0
567 1 943468 C C

Illustration by
Mike McQuade

FINDING SOMERTON MAN

How DNA, AI facial reconstruction, and sheer grit cracked a 75-year-old cold case • By Derek Abbott

The Somerton Man was photographed following the autopsy, in 1948 [top left]. The bump on his forehead is an artifact of the autopsy. He was buried by the Salvation Army on 14 June 1949 [bottom right]. The place on Somerton Beach where the man was found dead is marked with an X [bottom left]. Policemen recovered the man’s suitcase from the Adelaide city railway station and examined its contents [bottom middle]. The man’s fingerprints [top middle], taken after autopsy, were sent to the U.S. Federal Bureau of Investigation, which found no match.

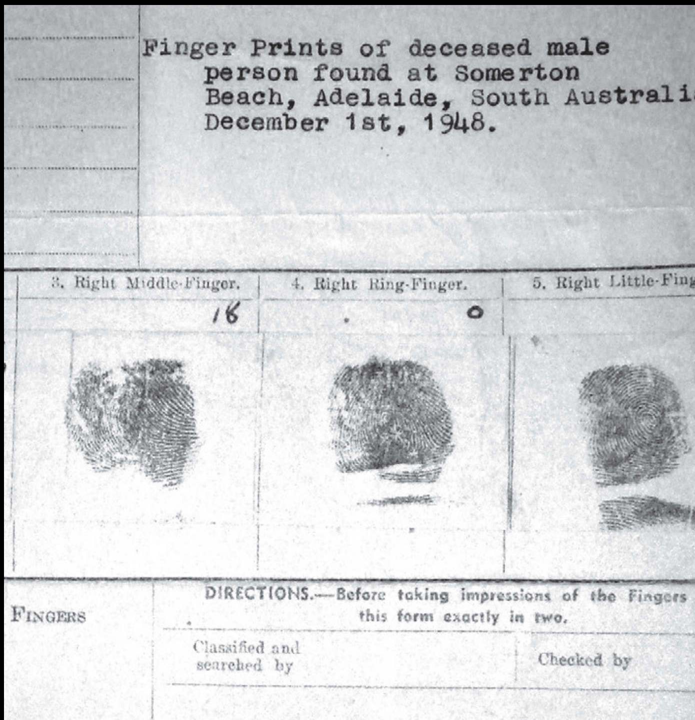
DEAD, AND IN A JACKET AND TIE. That’s how he was on 1 December 1948, when two men found him slumped against a retaining wall on the beach at Somerton, a suburb of Adelaide, Australia.

Police distributed a photograph, but no one came forward to claim the body. Eyewitnesses reported having seen the man, whom the newspapers dubbed the Somerton Man and who appeared to be in his early 40s, lying on the beach earlier, perhaps at one point moving his arm, and had concluded that he was drunk. The place of death led the police to treat the case as a suicide, despite the apparent lack of a suicide note. The presence of blood in the stomach, a common consequence of poisoning, was noted at the autopsy. Several chemical assays failed to identify any poison; granted, the methods of the day were not up to the task.

There was speculation of foul play. Perhaps the man was a spy who had come in from the cold; 1948 was the year after the Cold War got its name. This line of thought was strengthened, a few months later, by codelike writings in a book that came to be associated with the case.

These speculations aside, the idea that a person could simply die in plain view and without friends or family was shocking. This was a man with an athletic

CLOCKWISE FROM TOP LEFT: JAMES DURHAM; NEWS CORP./ALAMY (2)


Finger Prints of deceased male
person found at Somerton
Beach, Adelaide, South Australi[a]
December 1st, 1948.

3. Right Middle-Finger.	4. Right Ring-Finger.	5. Right Little-Fing[er]
16	. 0	

FINGERS — DIRECTIONS.—Before taking impressions of the Fingers this form exactly in two.

Classified and searched by — Checked by


Somerton body still mystery after 2 years

Tomorrow the Somerton body mystery will be two years old.

It is one of the most baffling cases the SA police have ever handled. The dead man's name and cause of his death are still unknown.

Early on the morning of December 1, 1948, Mr. John Lyons, of Whyte street, Somerton, found the body of a man leaning against the seawall opposite Somerton Crippled Children's Home. Name tags had been cut from his clothing.

The body was kept at the City Morgue for many weeks and was later embalmed by Mr. L. A. Elliott, of Hindmarsh. Scores of people saw it, many giving it different names.

(meaning "the end") was found on the body. These words are from "The Rubaiyat of Omar Khayyam."

The body was buried on June 14, 1949, at West Terrace Cemetery.

On June 22, 1949, the City Coroner (Mr. Cleland) adjourned the inquest to a date to be fixed in the hope that more evidence would be forthcoming.

He found the evidence before him was too inconclusive to warrant a finding.

Examinations by doctors and

CLOCKWISE FROM TOP LEFT: DEREK ABBOTT; TROVE PARTNERS; NEWS CORP./ALAMY

build, wearing a nice suit, and showing no signs of having suffered violence. The problem nagged many people over the years, and eventually it took hold of me. In the late 2000s, I began working on the Somerton Man mystery, devoting perhaps 10 hours a week to the research over the course of about 15 years.

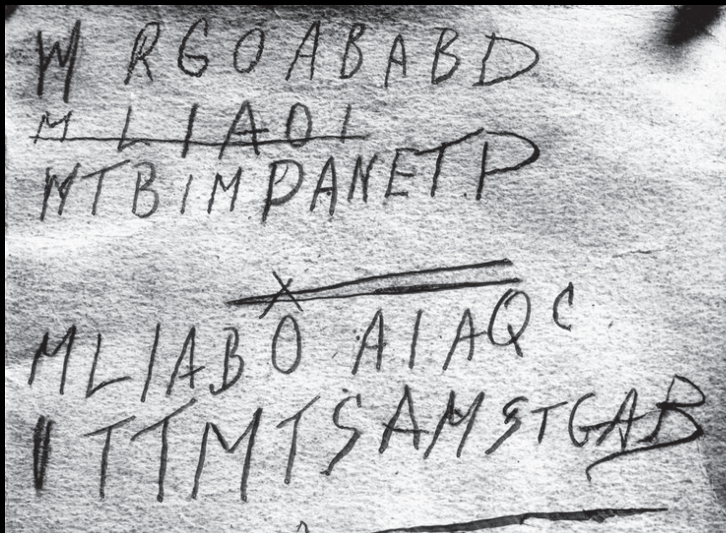
THROUGHOUT MY CAREER, I have always been interested in cracking mysteries. My students and I used computational linguistics to identify which of the three authors of *The Federalist Papers*—Alexander Hamilton, James Madison, and John Jay—was responsible for any given essay. We tried using the same method to confirm authorship of Biblical passages. More recently, we've been throwing some natural-language processing techniques into an effort to decode the Voynich Manuscript, an early 15th-century document written in an unknown language and an unknown script. These other projects yield to one or another key method of inquiry. The Somerton Man problem posed a broader challenge.

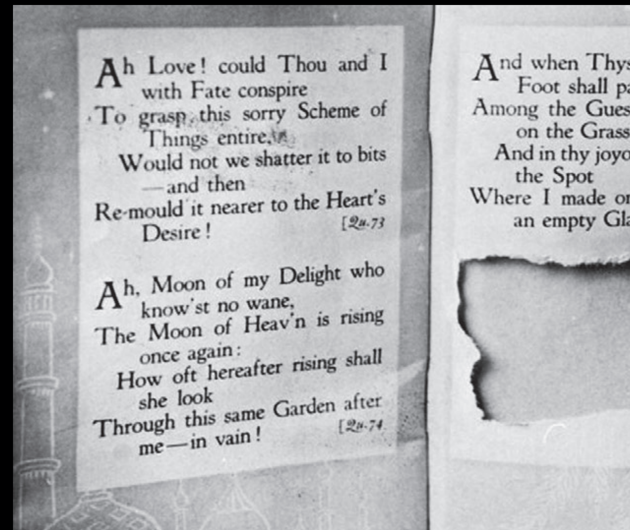
My one great advantage has been my access to students and to scientific instruments at the University of Adelaide, where I am a professor of electrical and electronic engineering. In 2009, I established a working group at the university's Center for Biomedical Engineering.

One question surrounding the Somerton Man had already been solved by sleuths of a more literary bent. In 1949, a pathologist had found a bit of paper concealed in one of the dead man's pockets, and on it were printed the words *Tamám Shud*, the Persian for "finished." The phrase appears at the end of Edward FitzGerald's translation of the *Rubáiyát of Omar Khayyám*, a poem that remains popular to this day.

The police asked the public for copies of the book in which the final page had been torn out. A man found such a book in his car, where apparently it had been thrown in through an open window. The book proved a match.

The back cover of the book also included scribbled letters, which were at first thought to constitute an encrypted message. But statistical tests carried out by my team showed that it was more likely a string of the initial letters of words. Through computational techniques, we eliminated all of the cryptographic codes known in the 1940s, leaving as a remaining possibility a one-time pad, in which each letter is based on a secret source text. We ransacked the poem itself and other texts, including the Bible and the Talmud, but



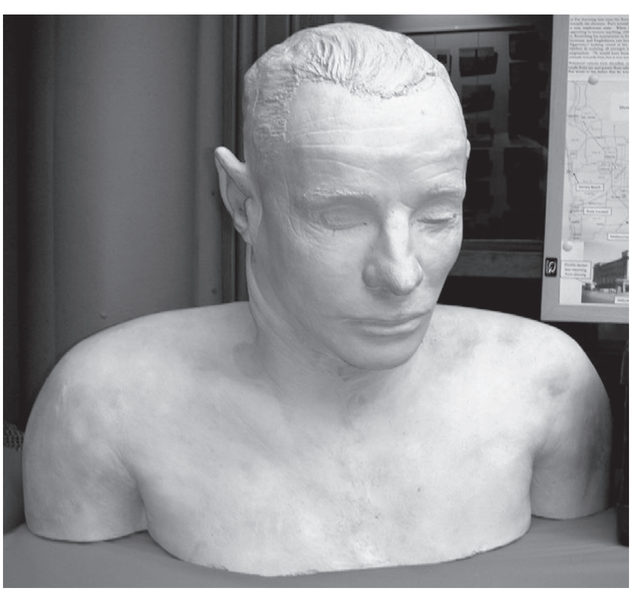


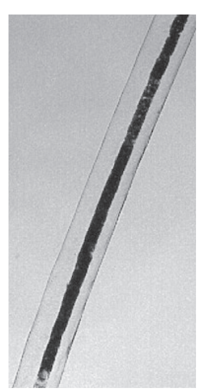
we never identified a plausible source text. It could have been a pedestrian *aide-mémoire*—to list the names of horses in an upcoming race, for example. Moreover, our research indicates that it doesn't have the structural sophistication of a code. The Persian phrase could have been the man's farewell to the world: his suicide note.

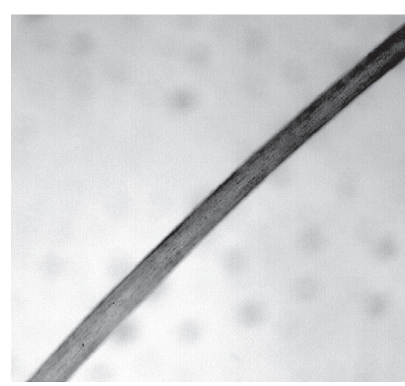
Also scribbled on the back cover was a telephone number that led to one Jo Thomson, a woman who lived merely a five-minute walk from where the Somerton Man had been found. Interviewers then and decades later reported that she had seemed evasive; after her death, some of her relatives and friends said they speculated that she must have known the dead man. I discovered a possible clue: Thomson's son was missing his lateral incisors, the two teeth that normally flank the central incisors. This condition, found in a very small percentage of the population, is often congenital; oddly, the Somerton Man had it, too. Were they related?

And yet the attempt to link Thomson to the body petered out. Early in the investigation, she told the police that she had given a copy of the *Rubáiyát* to a lieutenant in the Australian Army whom she had known during the war, and indeed, that man turned out to own a copy. But Thomson hadn't seen him since 1945, he was very much alive, and the last page of his copy was still intact. A trail to nowhere, one of many that were to follow.

WE ENGINEERS IN THE 21ST CENTURY had several other items to examine. First was a plaster death mask that had been made six months after the man died,

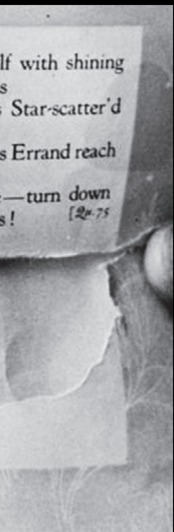


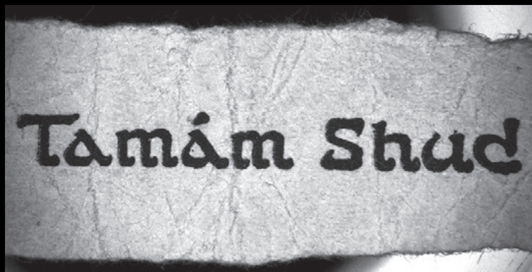




A plaster death mask [left] was molded directly from the cadaver 6 months after death, during which time the facial features had become distorted. The body was reported to have had graying hair at the sides; this gray hair [middle], shown under magnification, was pulled from the mask. A light brown hair was also found [right]; the man had been reported as having "mousey" colored hair.

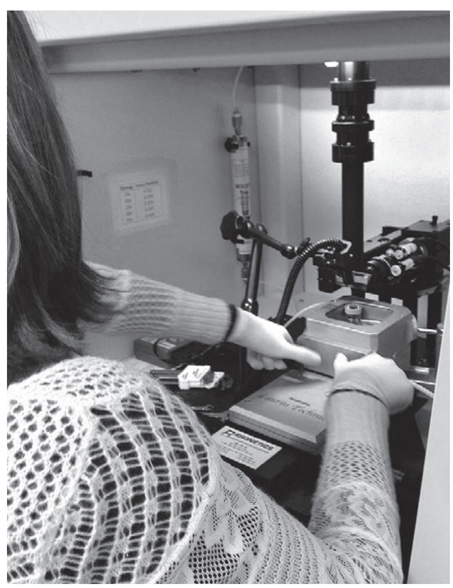
TOP, FROM LEFT: WILLIAM R.F. KRISCHOCK; NEWS CORP.; DEREK ABBOTT; T. GERALD KEANE

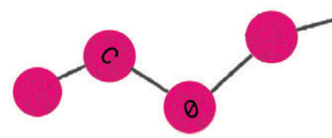


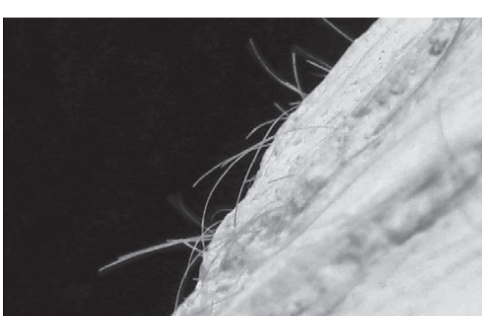


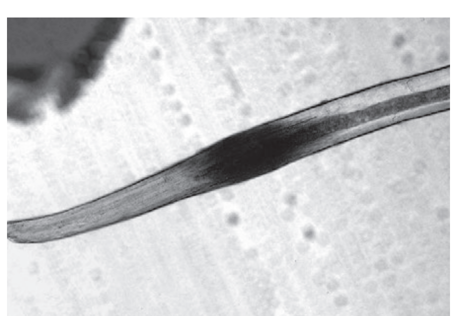
A copy of the *Rubáiyát of Omar Khayyám* was found months after the death of the Somerton Man. Letters scrawled on the back cover [top left] were at first mistaken for code. The book was linked to the dead man by a slip of paper [above] discovered in his watch pocket; the paper bore two Persian words. The slip corresponded to the missing part of the final page of the book [left]. The photograph [right] shows a youthful Charles Webb with his brother Roy and their parents, Dick and Eliza.

during which time the face had flattened. We tried several methods to reconstruct its original appearance: In 2013 we commissioned a picture by Greg O'Leary, a professional portrait artist. Then, in 2020, we approached Daniel Voshart, who designs graphics for Star Trek movies. He used a suite of professional AI tools to create a lifelike reconstruction of the Somerton Man. Later, we obtained another reconstruction by Michael Streed, a U.S. police sketch artist. We published these images, together with many isolated facts about



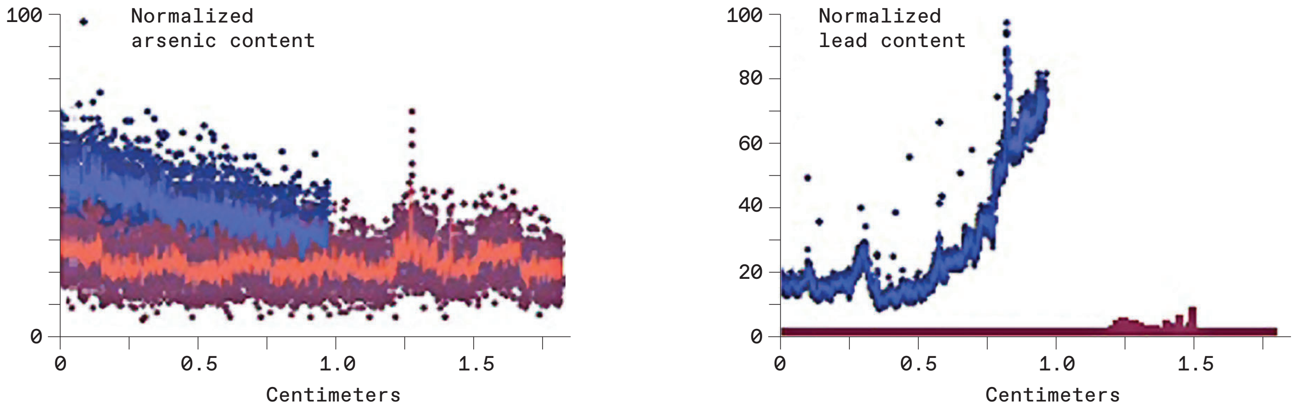






Clockwise from top left: Janette Edson is shown extracting hairs embedded in the plaster death mask. The hairs were then loaded into a laser ablation inductively coupled plasma mass spectrometer, which vaporizes hair along its shaft sequentially. A magnification of the Somerton Man's hair roots shows a darkened area—a phenomenon known as postmortem banding that can appear a few hours after death. Such banding is most pronounced in individuals who die slowly. This closeup view of the mask shows embedded hairs standing on end.

BOTTOM, CLOCKWISE FROM LEFT: DEREK ABBOTT (5); JANETTE EDSON & DEREK ABBOTT; DEREK ABBOTT



These two mass spectrographs compare hairs from the Somerton Man [blue] and from contemporary reference hair material [red]. The spectrograph for arsenic [left] shows an insignificantly elevated level in the Somerton Man. Note that the Somerton Man's hair sample is 1 centimeter long, which represents about a month's growth. The spectrograph for lead [right] dropped from a high level in the month before death, a clue that his routine changed during that time.

the body, the teeth, and the clothing, in the hope of garnering insights from the public. No luck.

As the death mask had been molded directly off the Somerton Man's head, neck, and upper body, some of the man's hair was embedded in the plaster of Paris—a potential DNA gold mine. At the University of Adelaide, I had the assistance of a hair forensics expert, Janette Edson. In 2012, with the permission of the police, Janette used a magnifying glass to find where several hairs came together in a cluster. She was then able to pull out single strands without breaking them or damaging the plaster matrix. She thus secured the soft, spongy hair roots as well as several lengths of hair shaft. The received wisdom of forensic science at the time held that the hair shaft would be useless for DNA analysis without the hair root.

Janette performed our first DNA analysis in 2015 and, from the hair root, was able to place the sample within a maternal genetic lineage, or haplotype, known as "H," which is widely spread around Europe. (Such maternally inherited DNA comes not from the nucleus of a cell but from the mitochondria.) The test therefore told us little we hadn't already known. The concentration of DNA was far too low for the technology of the time to piece together the sequencing we needed.

Fortunately, sequencing tools continued to improve. In 2018, Guanchen Li and Jeremy Austin, also at the University of Adelaide, obtained the entire mitochondrial genome from hair-root material and narrowed down the maternal haplotype to H4a1a1a.

However, to identify Somerton Man using DNA databases, we needed to go to autosomal DNA—the kind that is inherited from both parents. There are more than 20 such databases, 23andMe and Ancestry being the largest. These databases require sequences of from 500,000 to 2,000,000 single nucleotide polymorphisms, or SNPs (pronounced "snips"). The concentration levels of autosomes in the human cell tend to be much lower than those of the mitochondria, and so Li and Austin were able to obtain only 50,000 SNPs, of which 16,000 were usable. This was a breakthrough, but it wasn't good enough to work on a database.

IN 2022, AT THE SUGGESTION of Colleen Fitzpatrick, a former NASA employee who had trained as a nuclear physicist but then became a forensic genetics expert, I sent a hair sample to Astrea Forensics, a DNA lab in the United States. This was our best hair-root sample, one that I had nervously guarded for 10 years. The result from Astrea came back—and it was a big flop.

Seemingly out of options, we tried a desperate move. We asked Astrea to analyze a 5-centimeter-long shaft of hair that had no root at all. Bang! The company retrieved 2 million SNPs. The identity of the Somerton Man was now within our reach.

So why did the rootless shaft work in our case?

The DNA analysis that police use for standard crime-solving relies on only 20 to 25 short tandem repeats (STRs) of DNA. That's fine for police, who mostly do one-to-one matches to determine whether the DNA recovered at a crime scene matches a suspect's DNA.

But finding distant cousins of the Somerton Man on genealogical databases constitutes a one-to-many search, and for that you typically need around 500,000 markers. For these genealogical searches, SNPs are used because they contain information on ethnicity and ancestry generally. Note that SNPs have around 50 to 150 base pairs of nucleotides, whereas typical STRs tend to be longer, containing 80 to 450 base pairs. The hair shaft contains DNA that is mostly fragmented, so it's of little use when you're seeking longer STR segments but it's a great source of SNPs. So this is why crime forensics traditionally focused on the root and ignored the shaft, although this practice is now changing very slowly.

Another reason the shaft was such a trove of DNA is that keratin, its principal component, is a very tough protein, and it had protected the DNA fragments lodged

LEFT: JAMES CHAPPELL & DEREK ABBOTT; RIGHT: COLLEEN FITZPATRICK & DEREK ABBOTT within it. The 74-year-old soft spongy hair root, on the other hand, had not protected the DNA to the same extent. We set a world record for obtaining a human identification, using forensic genealogy, from the oldest piece of hair shaft. Several police departments in the United States now use hair shafts to retrieve DNA, as I am sure many will start to do in other countries, following our example.

Libraries of SNPs can be used to untangle the branching lines of descent in a family tree. We uploaded our 2 million SNPs to GEDmatch Pro, an online genealogical database located in Lake Worth, Fla. (and recently acquired by Qiagen, a biotech company based in the Netherlands). The closest match was a rather distant relative based in Victoria, Australia. Together with Colleen Fitzpatrick, I built out a family tree containing more than 4,000 people. On that tree we found a Charles Webb, son of a baker, born in 1905 in Melbourne, with no date of death recorded.

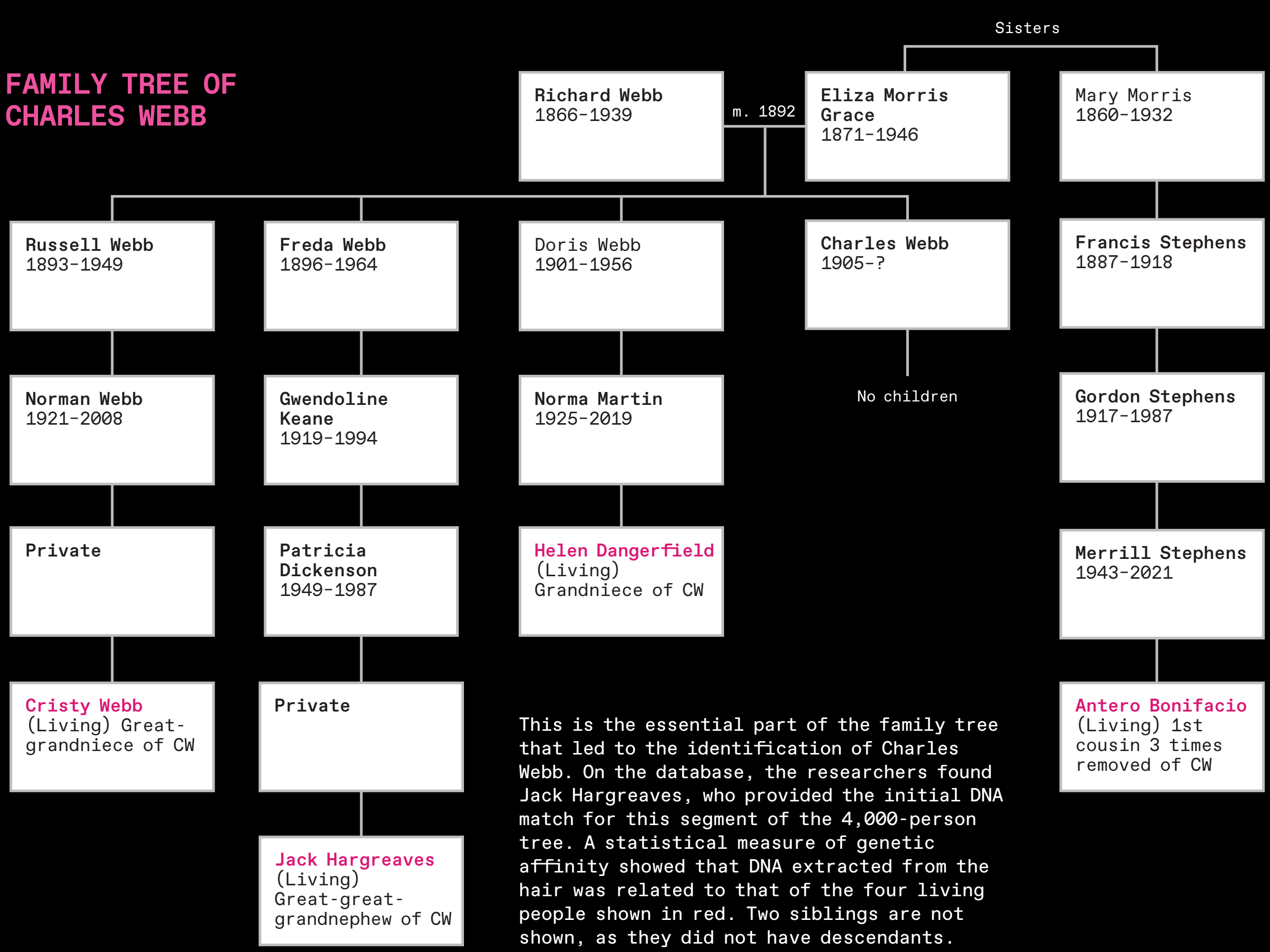
Charles never had children of his own, but he had five siblings, and I was able to locate some of their living descendants. Their DNA was a dead match. I also found a descendant of one of his maternal aunts, who agreed to undergo a test. When a positive result came through on 22 July 2022, we had all the evidence we needed. This was our champagne moment.

IN LATE 2021, POLICE in South Australia ordered an exhumation of the Somerton Man's body for a thorough analysis of its DNA. At the time we prepared this article, they had not yet confirmed our result, but they did announce that they were "cautiously optimistic" about it.

All at once, we were able to fill in a lot of blank spaces. Webb was born on 16 November 1905, in Footscray, a suburb of Melbourne, and educated at a technical college, now Swinburne University of Technology. He later worked as an electrical technician at a factory that made electric hand drills. Our DNA tests confirmed he was not related to Thomson's son, despite the coincidence of their missing lateral incisors.

We discovered that Webb had married a woman named Dorothy Robertson in 1941 and had separated from her in 1947. She filed for divorce on grounds of desertion, and the divorce lawyers visited his former place of work, confirming that he had quit around 1947 or 1948. But they could not determine what happened to him after that. The divorce finally came through

FAMILY TREE OF CHARLES WEBB



This is the essential part of the family tree that led to the identification of Charles Webb. On the database, the researchers found Jack Hargreaves, who provided the initial DNA match for this segment of the 4,000-person tree. A statistical measure of genetic affinity showed that DNA extracted from the hair was related to that of the four living people shown in red. Two siblings are not shown, as they did not have descendants.

Daniel Voshart used AI-based software to produce this reconstruction of Charles Webb's appearance.

DANIEL VOSHART

in 1952; in those days, divorces in Australia were granted only five years after separation.

At the time of Webb's death his family had become quite fragmented. His parents were dead, a brother and a nephew had died in the war, and his eldest brother was ill. One of his sisters died in 1955 and left him money in her will, mistakenly thinking he was still alive and living in another state. The lawyers administering the will were unable to locate Charles.

We got more than DNA from the hair: We also vaporized a strand of hair by scanning a laser along its length, a technique known as laser ablation. By performing mass spectrometry on the vapor, we were able to track Webb's varying exposure to lead. A month before Webb's death, his lead level was high, perhaps because he had been working with the metal, maybe soldering with it. Over the next month's worth of hair growth, the lead concentration declined; it reached its lowest level at his death. This might be a sign that he had moved.

With a trove of photographs from family albums and other sources, we were able to compare the face of the young Webb with the artists' reconstructions we had commissioned in 2013 and 2021 and the AI reconstruction we had commissioned in 2020. Interestingly, the AI reconstruction had best captured his likeness.

A group photograph, taken in 1921, of the Swinburne College football team, included a young Webb. Clues found in newspapers show that he continued to participate in various sports, which would explain the athletic condition of his body.

What's interesting about solving such a case is how it relies on concepts that may seem counterintuitive to forensic biologists but are quite straightforward to an electronics engineer. For example, when dealing with a standard crime scene that uses only two dozen STR markers, one observes very strict protocols to ensure the integrity of the full set of STRs. When dealing with a case with 2 million SNPs, by contrast, things are more relaxed. Many of the old-school STR protocols don't apply when you have access to a lot of information. Many SNPs can drop out, some can even be "noise," the signal may not be clean—and yet you can still crack the case!

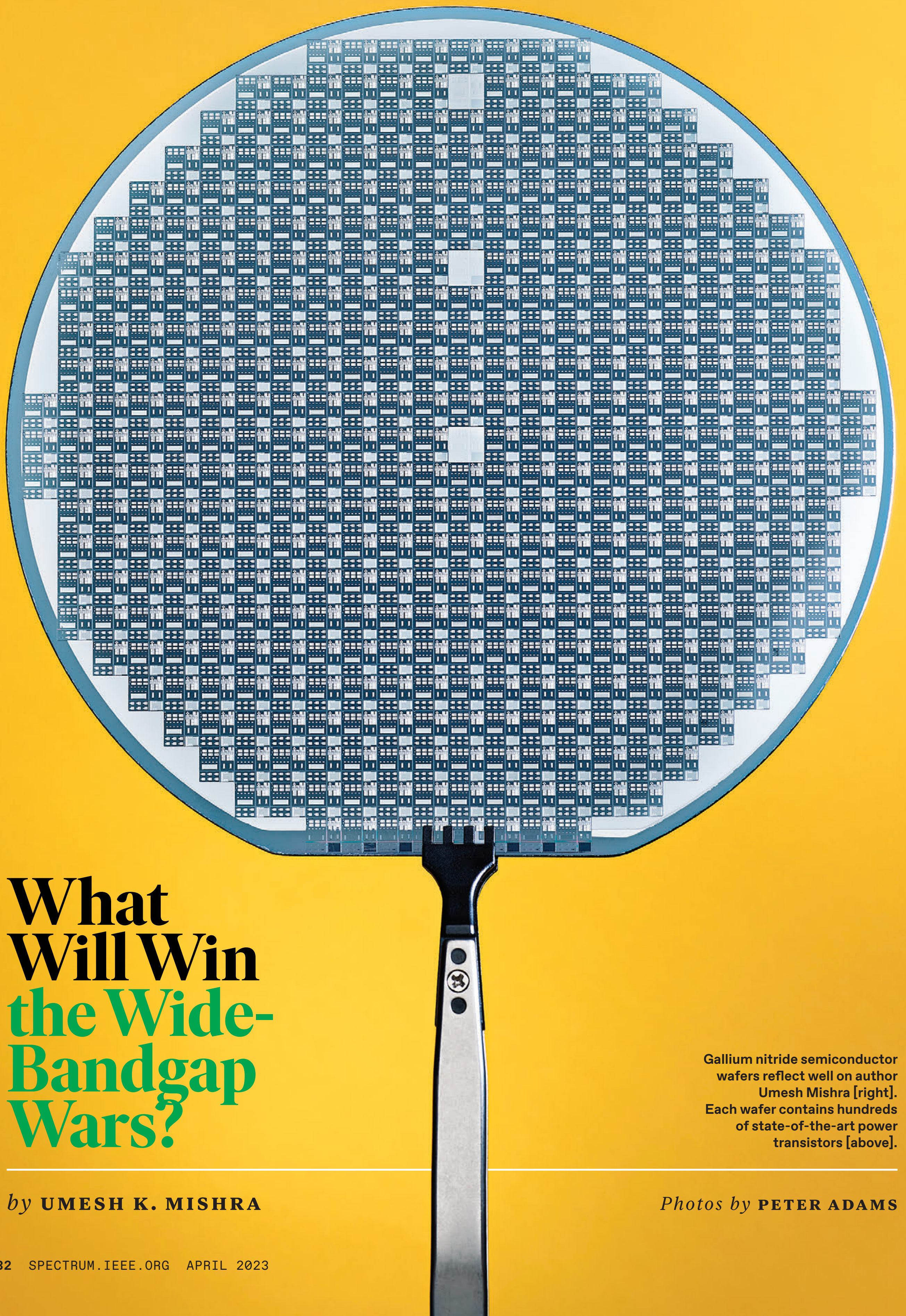
Engineers understand this concept well. It's what we call graceful degradation—when, say, a few flipped bits on a digital video signal are hardly noticed. The same is true for a large SNP file.

And so, when Astrea retrieved the 2 million SNPs, the company didn't rely on the traditional framework for DNA-sequencing reads. It used a completely different mathematical framework, called imputation. The concept of imputation is not yet fully appreciated by forensics experts who have a biological background. However, for an electronics engineer, the concept is similar to error correction: We infer and "impute" bits of information that have dropped out of a received digital signal. Such an approach is not possible with a few STRs, but when handling over a million SNPs, it's a different ball game.

Much of the work on identifying Charles Webb from his genealogy had to be done manually because there are simply no automated tools for the task. As an electronics engineer, I now see possible ways to make tools that would speed up the process. One such tool my team has been working on, together with Colleen Fitzpatrick, is software that can input an entire family tree and represent all of the birth locations as colored dots on Google Earth. This helps to visualize geolocation when dealing with a large and complex family.

The Somerton Man case still has its mysteries. We cannot yet determine where Webb lived in his final weeks or what he was doing. Although the literary clue he left in his pocket was probably an elliptical suicide note, we cannot confirm the exact cause of death. There is still room for research; there is much we do not know. ■


IEEE Women in Engineering
Wie
Connect. Support. Inspire.



What Will Win the Wide-Bandgap Wars?

Gallium nitride semiconductor wafers reflect well on author Umesh Mishra [right]. Each wafer contains hundreds of state-of-the-art power transistors [above].

by UMESH K. MISHRA

Photos by PETER ADAMS

As **gallium nitride** and **silicon carbide** vie for dominance, they will cut greenhouse gases by billions of tonnes

Can advanced semiconductors cut emissions of greenhouse gases enough to make a difference in the struggle to halt climate change?

The answer is a resounding yes. Such a change is actually well underway.

Starting around 2001, the compound semiconductor gallium nitride fomented a revolution in lighting that has been, by some measures, the fastest technology shift in human history. In just two decades, the share of the global lighting market held by gallium-nitride-based light-emitting diodes has gone from zero to more than 50 percent, according to a study by the International Energy Agency. The research firm Mordor Intelligence recently predicted that, worldwide, LED lighting will be responsible for cutting the electricity used for lighting by 30 to 40 percent over the next seven years. Globally, lighting accounts for about 20 percent of electricity use and 6 percent of carbon dioxide emissions, according to the United Nations Environment Program.

This revolution is nowhere near done. Indeed, it is about to jump to a higher level. The very semiconductor technology that has transformed the lighting industry, gallium nitride (GaN), is also part of a revolution in power electronics that is now gathering steam. It is one of two semiconductors—the other being silicon carbide (SiC)—that have begun displacing silicon-based electronics in enormous and vital categories of power electronics.

GaN and SiC devices perform better and are more efficient than the silicon components they are replacing. There are countless billions of these devices all over the world, and many of them operate for hours every day, so the energy savings are going to be substantial. The rise of GaN and SiC power electronics will ultimately have a greater positive impact on the planet's climate than will the replacement of incandescent and other legacy lighting by GaN LEDs.

Virtually everywhere that alternating current must be transformed to direct current or vice versa, there will be fewer wasted watts. This conversion happens in your phone's or laptop's wall charger, in the much larger chargers and inverters that power electric vehicles, and elsewhere. And there will be similar savings as other silicon strongholds fall to the new semiconductors, too. Wireless base-station amplifiers are among the growing applications for which these emerging semiconductors are clearly superior. In the effort to mitigate climate change, eliminating waste in power consumption is the low-hanging fruit, and these semiconductors are the way we'll harvest it.

This is a new instance of a familiar pattern in technology history: two competing innovations coming to fruition at the same time. How will it all shake out? In which applications will SiC dominate, and in which will GaN prevail? A hard look at the relative strengths of these two semiconductors gives us some solid clues.

BEFORE WE GET TO the semiconductors themselves, let's first consider why we need them. To begin with: Power conversion is everywhere. And it goes far beyond the little wall chargers that sustain our smartphones, tablets, laptops, and countless other gadgets.

Power conversion is the process that changes electricity from the form that's available to the form required for a product to perform its function. Some energy is always lost in that conversion, and because some of these products run continuously, the energy savings can be enormous. Consider: Electricity consumption in the state of California remained essentially flat from 1980 even as the economic output of the state skyrocketed. One of the most important reasons why the demand remained flat is that the efficiency of refrigerators and air conditioners increased enormously over that period. The single-greatest factor in this improvement has been the use of variable-speed drives based on the insulated gate bipolar transistor (IGBT) and other power electronics, which greatly increased efficiency.

SiC and GaN are going to enable far greater reductions in emissions. GaN-based technologies alone could lead to a savings of over 1 billion tonnes of greenhouse gases in 2041 in just the United States and India, according to an analysis of publicly available data by Transphorm, a GaN-device company I cofounded in 2007. The data came from the International Energy Agency, Statista, and other sources. The same analysis indicates a 1,400-terawatt-hour energy savings—or 10 to 15 percent of the projected energy consumption by the two countries that year.

LIKE AN ORDINARY TRANSISTOR, a power transistor can act as an amplifying device or as a switch. An important example of the amplifying role is in wireless base stations, which amplify signals for transmission to smartphones. All over the world, the semiconductor used to fabricate the transistors in these amplifiers is shifting from a silicon technology called laterally diffused metal-oxide semiconductor (LDMOS) to GaN. The newer technology has many advantages, including a power-efficiency improvement of 10 percent or more depending on frequencies. In power-conversion applications, on the other hand, the transistor acts as a switch rather than as an amplifier. The standard technique is called pulse-width modulation. In a common type of motor controller, for example, pulses of direct-current electricity are fed to coils mounted on the motor's rotor. These pulses set up a magnetic field that interacts with that of the motor's stator, which makes the rotor spin. The speed of this rotation is controlled by altering the length of the pulses: A graph of these pulses is a square wave, and the longer the pulses are "on" rather than "off," the more rotational speed and torque the motor provides. Power transistors accomplish the on-and-off switching.

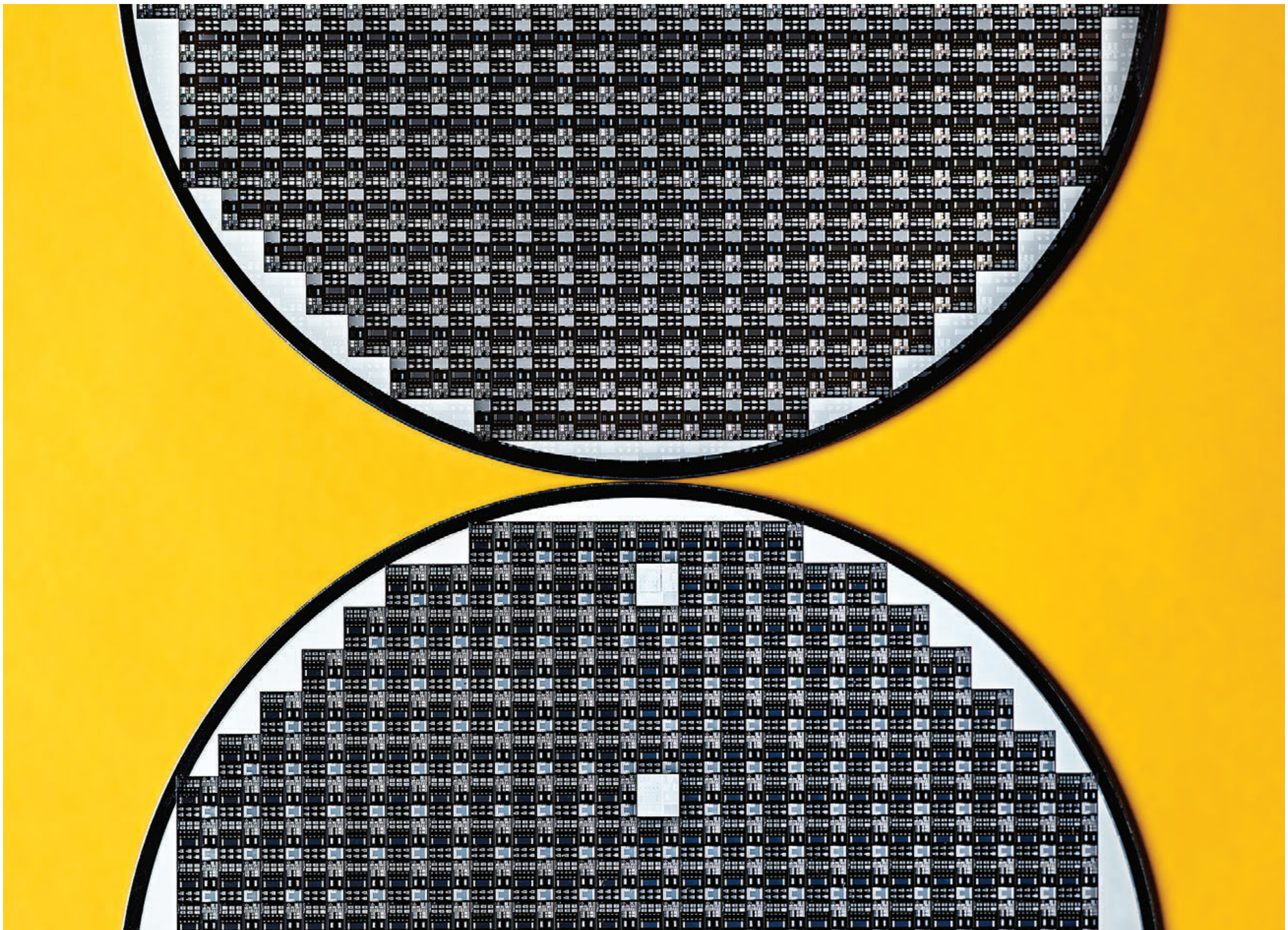
Pulse-width modulation is also used in switching power supplies, one of the most common examples of power conversion. Switching power supplies are the type used to power virtually all personal computers, mobile devices, and appliances that run on DC. Basically, the input AC voltage is converted to DC, and then that DC is "chopped" into a high-frequency alternating-current square wave. This chopping is done by power transistors, which create the square wave by switching the DC on and off. The square wave is applied to a transformer that changes the amplitude of the wave to produce the desired output voltage. To get a steady DC output, the voltage from the transformer is rectified and filtered.

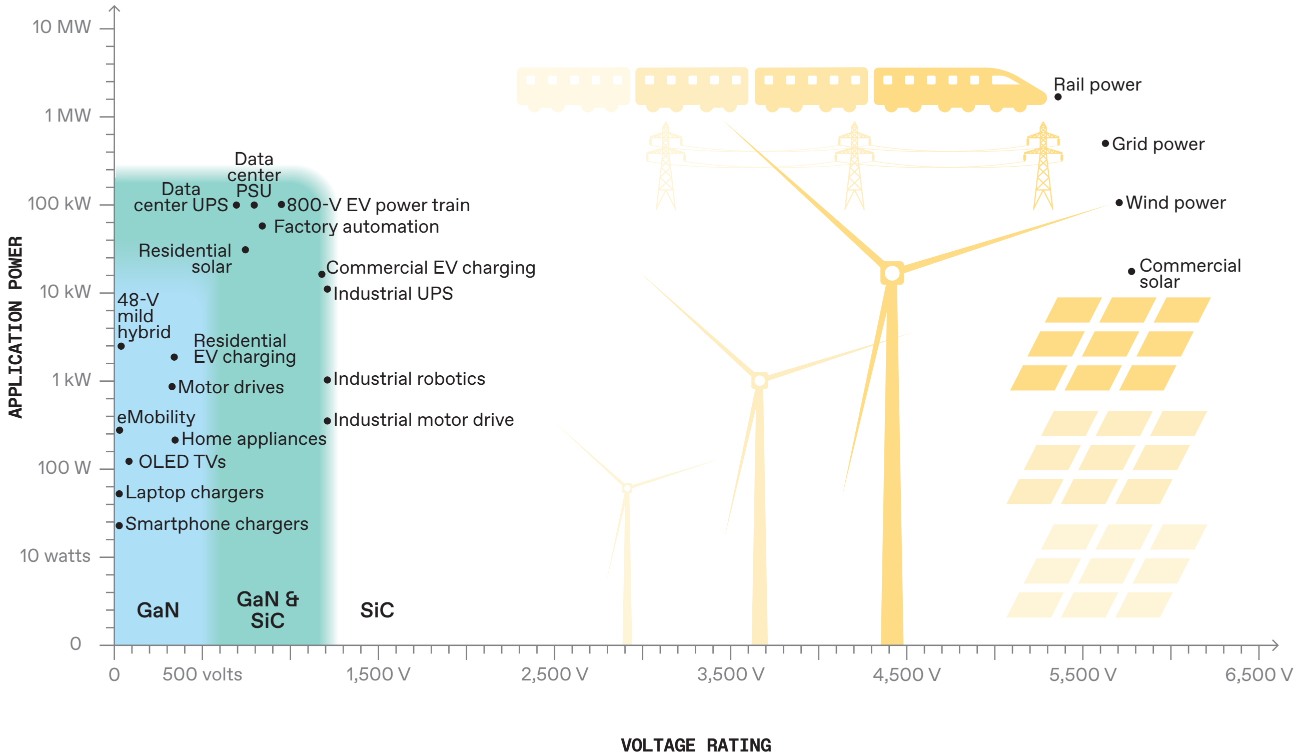
The important point here is that the characteristics of the power transistors determine, almost entirely, how well the circuits can perform pulse-width modulation—and therefore, how efficiently the controller regulates the voltage. An ideal power transistor would, when in the off state, completely block current flow even when the applied voltage is high. This characteristic is called high electric breakdown field strength, and it indicates how much voltage the semiconductor is able to withstand. On the other hand, when it is in the on state, this ideal transistor would have very low resistance to the flow of current. This feature results from very high mobility of the charges—electrons and holes—within the semiconductor's crystalline lattice. Think of breakdown field strength and charge mobility as the yin and yang of a power semiconductor.

GaN and SiC come much closer to this ideal than the silicon semiconductors they are replacing. First, consider breakdown field strength. Both GaN and SiC belong to a class called wide-bandgap semiconductors. The bandgap of a semiconductor is defined as the energy, in electron volts, needed for an electron in the semiconductor lattice to jump from the valence band to the conduction band. An electron in the valence band participates in the bonding of atoms within the crystal lattice, whereas in the conduction band electrons are free to move around in the lattice and conduct electricity.

In a semiconductor with a wide bandgap, the bonds between atoms are strong and so the material is usually able to withstand relatively high voltages before the bonds break and the transistor is said to break down. The bandgap of silicon is 1.12 electron volts, as compared with 3.40 eV for GaN. For the most common type of SiC, the band gap is 3.26 eV.

Now let's look at mobility, which is given in units of centimeters squared per volt second ($cm^2/V{\cdot}s$). The product of mobility and electric field yields the velocity of the electron,





GALLIUM NITRIDE AND SILICON CARBIDE: WHERE THEY COMPETE

In the markets for high-voltage power transistors, gallium nitride devices dominate in applications below around 500 volts [blue in chart above], while silicon carbide has the edge now for 1000 V and above (the markets are relatively small above around 2,000 V). The landscape of the important battleground between 500 and 1,200 V [green] will change as GaN devices improve. For example, with the introduction of 1,200-V GaN transistors—expected in 2025—the battle will be joined in the all-important market for electric-vehicle inverters.

and the higher the velocity the higher the current carried for a given amount of moving charge. For silicon this figure is 1,450; for SiC it is around 950; and for GaN, about 2,000. GaN's unusually high value is the reason why it can be used not only in power-conversion applications but also in microwave amplifiers. GaN transistors can amplify signals with frequencies as high as 100 gigahertz—far above the 3 to 4 GHz generally regarded as the maximum for silicon LDMOS. For reference, 5G's millimeter-wave frequencies top out at 52.6 GHz. This highest 5G band is not yet widely used; however, frequencies up to 75 GHz are being deployed in dish-to-dish communications, and researchers are now working with frequencies as high as 140 GHz for in-room communications. The appetite for bandwidth is insatiable.

These performance figures are important, but they're not the only criteria by which GaN and SiC should be compared for any particular application. Other critical factors include ease of use and cost, for both the devices and the systems into which they are integrated. Taken together, these factors explain where and why each of these semiconductors has begun displacing silicon. [See illustration, "Gallium Nitride and Silicon Carbide: Where They Compete."] They also offer strong clues about how their future competition may shake out.

THE FIRST COMMERCIALLY VIABLE SiC transistor that was superior to silicon was introduced by Cree (now Wolfspeed) in 2011. It could block 1,200 volts and had a respectably low resistance of 80 milliohms when conducting current. Today there are three different kinds of SiC transistors on the market. There's a trench MOSFET (metal-oxide semiconductor field-effect transistor) from Rohm; DMOSs (double-diffused MOSs) from Infineon Technologies, ON Semiconductor Corp., STMicroelectronics, Wolfspeed, and others; and a vertical-junction field-effect transistor from Qorvo.

One of the big advantages of SiC MOSFETs is their similarity to traditional silicon ones—even the packaging is identical. A SiC MOSFET operates in essentially the same way as an ordinary silicon MOSFET. There's a source, a gate, and a drain. When the device is on, electrons flow from a heavily doped *n*-type source across a lightly doped bulk region before being "drained" through a conductive substrate. This similarity means that there's a minimal learning curve for engineers making the switch to SiC.

Compared to GaN, SiC has other advantages. SiC MOSFETs are inherently "fail-open" devices, meaning that if the control circuit fails for any reason the transistor stops conducting current. This is an important feature, because this characteristic largely eliminates the possibility that a failure could lead to a short circuit and a fire or explosion. (The price paid for this feature, however, is a lower electron mobility, which increases resistance when the device is on.)

Illustrations by Chris Philpot

GaN brings its own unique advantages. The semiconductor first established itself commercially in 2000 in the markets for light-emitting diodes and semiconductor lasers. It was the first semiconductor capable of reliably emitting bright green, blue, purple, and ultraviolet light. But long before this commercial breakthrough in optoelectronics, I and other researchers had already demonstrated the promise of GaN for high-power electronics. GaN LEDs caught on quickly because they filled a void for efficient lighting. But GaN for electronics had to prove itself superior to existing technologies: in particular, silicon CoolMOS transistors from Infineon for power electronics, and silicon-LDMOS and gallium-arsenide transistors for radio-frequency electronics.

GaN's main advantage is its extremely high electron mobility. Electric current, the flow of charge, equals the concentration of the charges multiplied by their velocity. So you can get high current because of high concentration or high velocity or some combination of the two. The GaN transistor is unusual because most of the current flowing through the device is due to electron velocity rather than charge concentration. What this means in practice is that, in comparison with silicon or SiC, less charge has to flow into the device to switch it on or off. That, in turn, reduces the energy needed for each switching cycle and contributes to high efficiency.

Meanwhile, GaN's high electron mobility allows switching speeds on the order of 50 V per nanosecond. That characteristic means power converters based on GaN transistors operate efficiently at frequencies in the multiple hundreds of kilohertz, as opposed to about 100 kilohertz for silicon or SiC.

Taken together, the high efficiency and high frequency enables the power converter based on GaN devices to be quite small and lightweight: High efficiency means smaller heat sinks, and operation at high frequencies means that the inductors and capacitors can be very small, too.

One disadvantage of GaN semiconductors is that they do not yet have a reliable insulator technology. This complicates the design of devices that are fail-safe—in other words, that fail open if the control circuit fails.

There are two options to achieve this normally off characteristic. One is to equip the transistor with a type of gate that removes the charge in the channel when there's no voltage applied to the gate and that conducts current only on application of a positive voltage to that gate. These are called enhancement-mode devices. They are offered by EPC, GaN Systems, Infineon, Innoscience, and Navitas, for example. [See illustration, "Enhancement-Mode GaN Transistor."]

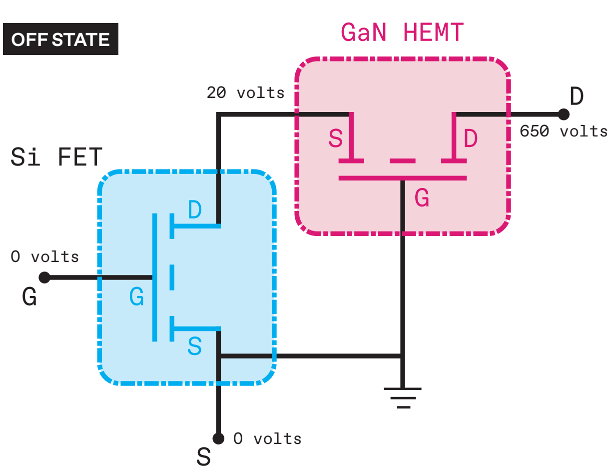
The other option is called the cascode solution. It uses a separate, low-loss silicon field-effect transistor to provide the fail-safe feature for the GaN transistor. This cascode solution is used by Power Integrations, Texas Instruments, and Transphorm. [See illustration, "Cascoded Depletion-Mode GaN Transistor."]

No comparison of semiconductors is complete without a consideration of costs. A rough rule of thumb is—smaller die size means lower cost. Die size is the physical area of the integrated circuit containing the devices.

SiC devices now generally have smaller dies than GaN ones. However, SiC's substrate and fabrication costs are higher than those for GaN and, in general, the final device costs for applications at 5 kilowatts and higher are not much different today. Future trends, though, are likely to favor GaN. I base this belief on the relative simplicity of GaN devices, which will mean production costs low enough to overcome the larger die size.

CASCODED DEPLETION-MODE GaN TRANSISTOR

For safety, when a power transistor's control circuit fails, it must fail into the open state, with no current flow. This is a challenge for gallium nitride devices because they lack a gate-insulator material that is reliable both in the high-voltage blocking state and in the current-carrying on state. One solution, called cascoded depletion mode, uses a low-voltage signal on a silicon field-effect transistor (FET) to control the much larger voltage on a gallium nitride high electron mobility transistor [right]. If the control circuit fails, the voltage on the gate of the FET drops to zero and it stops conducting current [left]. With the FET no longer conducting current, the gallium nitride transistor also stops conducting, because there is no longer a closed circuit between the drain and the source of the combined device.



That said, for GaN to be viable for many high-power applications that also demand high voltages, it must have a cost-effective, high-performance device rated for 1,200 V. After all, there are already SiC transistors available at that voltage. Currently, the closest commercially available GaN transistors are rated for 900 V, produced by Transphorm, based in Goleta, Calif., which I cofounded with Primit Parikh. Lately, we have also demonstrated 1,200-V devices, fabricated on sapphire substrates, that have both electrical and thermal performance on a par with SiC devices.

Projections from the research firm Omdia for 1,200-V SiC MOSFETs indicate a price of 16 cents per ampere in 2025. In my estimation, because of the lower cost of GaN substrates, the price of first-generation 1,200-V GaN transistors in 2025 will be less than that of their SiC counterparts. Of course, that's just my opinion; we'll all know for sure how this will shake out in a couple of years.

WITH THESE RELATIVE ADVANTAGES and disadvantages in mind, let's consider individual applications, one by one, and shed some light on how things might develop.

• ELECTRIC VEHICLE INVERTERS AND CONVERTERS: Tesla's adoption of SiC in 2017 for the onboard, or traction, inverters for its Model 3 was an early and major win for the semiconductor. In an EV, the traction inverter converts the DC from the batteries to AC for the motor. The inverter also controls the speed of the motor by varying the frequency of the alternating current. Today, Mercedes-Benz and Lucid Motors are also using SiC in their inverters, and other EV makers are planning to use SiC in upcoming models, according to news reports. The SiC devices are being supplied by Infineon, OnSemi, Rohm, Wolfspeed, and others. EV traction inverters typically range from about 35 kW to 100 kW for a small EV to about 400 kW for a large vehicle.

However, it's too soon to call this contest for SiC. As I noted, to make inroads in this market, GaN suppliers will have to offer a 1,200-V device. EV electrical systems now typically operate at just 400 V, but the Porsche Taycan has an 800-V system, as do EVs from Audi, Hyundai, and Kia. Other automakers are expected to follow their lead in coming years. (The Lucid Air has a 900-V system.) I expect to see the first commercial 1,200-V GaN transistors in 2025. These devices will be used not only in vehicles but also in high-speed public EV chargers.

The higher switching speeds possible with GaN will be a powerful advantage in EV inverters, because these switches employ what are called hard-switched techniques. Here, the way to enhance performance is to switch very fast from on to off to minimize the time when the device is both holding high voltage *and* passing high current.

Besides an inverter, an EV also typically has an **onboard charger**, which enables the vehicle to be charged from wall (mains) current by converting AC to DC. Here, again, GaN is very attractive, for the same reasons that make it a good choice for inverters.

• ELECTRIC-GRID APPLICATIONS: Very-high-voltage power conversion for devices rated at 3 kilovolts and higher will remain the domain of SiC for at least the next decade. These applications include systems to help stabilize the grid, convert AC to DC and back again at transmission-level voltages, and other uses.

• PHONE, TABLET, AND LAPTOP CHARGERS: Starting in 2019, GaN-based wall chargers became available commercially from companies such as GaN Systems, Innoscience, Navitas, Power Integrations, and Transphorm. The high switching

ENHANCEMENT-MODE GaN TRANSISTOR

One of the two major types of gallium nitride transistor is called an enhancement-mode device. It uses a gate-control circuit operating at around 6 volts to control the main switching circuit, which can block 600 V or more when the control circuit is off. When the device is on (when 6 V are applied to the gate), electrons flow from the drain to the source in a flat region called a two-dimensional electron gas. In this region the electrons are extremely mobile—a factor that helps enable very high switching speeds—and confined beneath a barrier of aluminum gallium nitride. When the device is off, the region below the gate is depleted of electrons, breaking the circuit under the gate and stopping current flow.

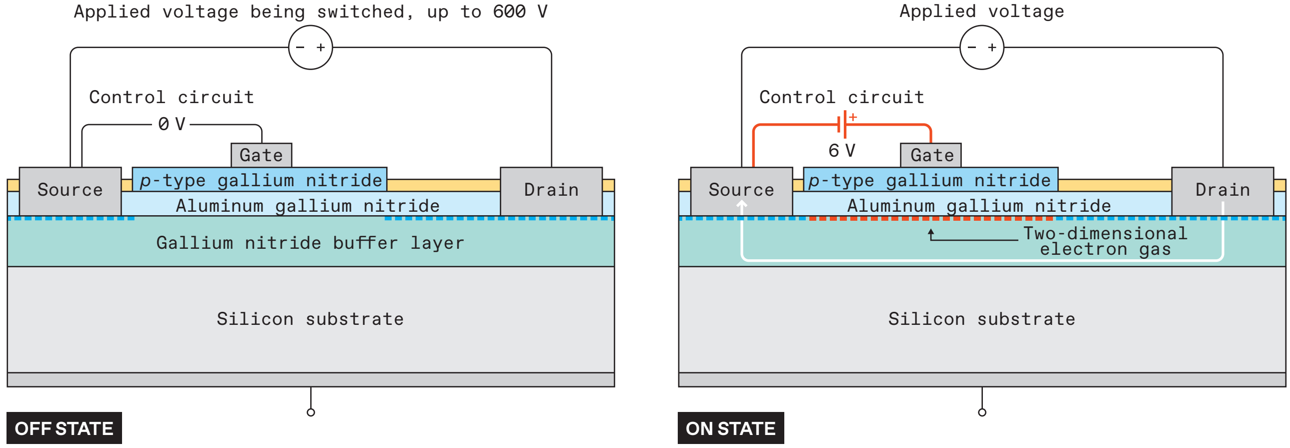
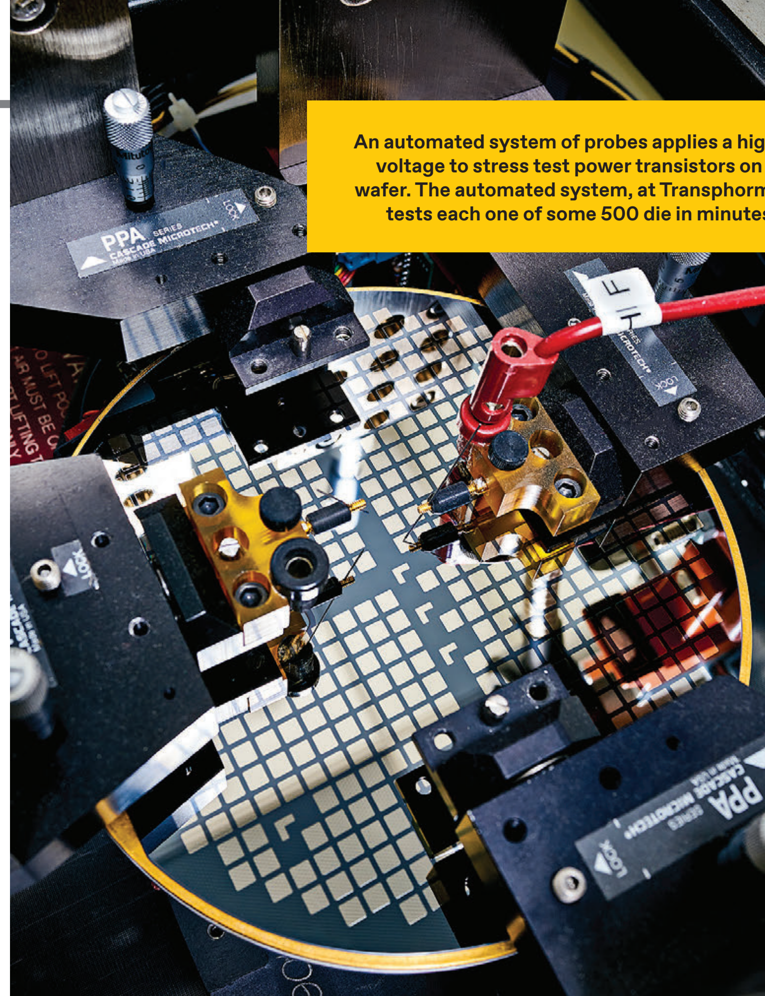


Illustration by Chris Philpot

speeds of GaN coupled with its generally lower costs have made it the incumbent in lower-power markets (25 to 500 W), where these factors, along with small size and a robust supply chain, are paramount. These early GaN power converters had switching frequencies as high as 300 kHz and efficiencies above 92 percent. They set records for power density, with figures as high as 30 W per cubic inch (1.83 W/cm^3)—roughly double the density of the silicon-based chargers they are replacing.

• **SOLAR-POWER MICROINVERTERS:** Solar-power generation has taken off in recent years, in both grid-scale and distributed (household) applications. For every installation, an inverter is needed to convert the DC from the solar panels to AC to power a home or release the electricity to the grid. Today, grid-scale photovoltaic inverters are the domain of silicon IGBTs and SiC MOSFETs. But GaN will begin making inroads in the distributed solar market, particularly.

Traditionally, in these distributed installations, there was a single inverter box for all of the solar panels. But increasingly installers are favoring systems in which there is a separate microinverter for each panel, and the AC is combined before powering the house or feeding the grid. Such a setup means the system can monitor the operation of each panel in order to optimize the performance of the whole array.



An automated system of probes applies a high voltage to stress test power transistors on a wafer. The automated system, at Transphorm, tests each one of some 500 die in minutes.

Microinverter or traditional inverter systems are critical to the modern data center. Coupled with batteries they create an uninterruptible power supply to prevent outages. Also, all data centers use power-factor correction circuits, which adjust the power supply's alternating-current waveforms to improve efficiency and remove characteristics that could damage equipment. And for these, GaN provides a low-loss and economical solution that is slowly displacing silicon.

• **5G AND 6G BASE STATIONS:** GaN's superior speed and high power density will enable it to win and ultimately dominate applications in the microwave regimes, notably 5G and 6G wireless, and commercial and military radar. The main competition here are arrays of silicon LDMOS devices, which are cheaper but have lower performance. Indeed, GaN has no real competitor at frequencies of 4 GHz and above.

For 5G and 6G wireless, the critical parameter is bandwidth, because it determines how much information the hardware can transmit efficiently. Next-generation 5G systems will have nearly 1 GHz of bandwidth, enabling blazingly fast video and other applications.

Microwave-communication systems that use silicon-on-insulator technologies provide a 5G+ solution using high-frequency silicon devices where each device's low output power is overcome with large arrays of them. GaN and silicon will coexist for a while in this space. The winner in a specific application will be determined by a tradeoff among system architecture, cost, and performance.

• **RADAR:** The U.S. military is deploying many ground-based radar systems based on GaN electronics. These include the Ground/Air Task Oriented Radar and the Active Electronically Scanned Array Radar built by Northrup-Grumman for the U.S. Marine Corps. Raytheon's SPY6 radar was delivered to the U.S. Navy and tested for the first time at sea in December 2022. The system greatly extends the range and sensitivity of shipborne radar.

TODAY, SIC DOMINATES in EV inverters, and generally wherever voltage-blocking capability and power handling are paramount and where the frequency is low. GaN is the preferred technology where high-frequency performance matters, such as in base stations for 5G and 6G, and for radar and high-frequency power-conversion applications such as wall-plug adapters, microinverters, and power supplies.

But the tug-of-war between GaN and SiC is just beginning. Regardless of how the competition plays out, application by application and market by market, we can say for sure that the Earth's environment will be a winner. Countless billions of tonnes of greenhouse gases will be avoided in coming years as this new cycle of technological replacement and rejuvenation wends its way inexorably forward. ■

This article was jointly produced by *IEEE Spectrum* and *Proceedings of the IEEE* with similar versions published in both publications.

The EV Transition Is Harder Than Anyone Thinks

Clueless policymakers, skeptical consumers, greedy automakers—and the tech isn't ready either **By Robert N. Charette**

Illustration by Christian Gralingen


8
6
23
15
6
R30
R25
23
40
R20
R10
25
18
15
50
45°
96
70
55
96
R10
70
140

Volvo Cars CEO Jim Rowan boldly proclaims that electric vehicles will reach price parity with internal-combustion-engine (ICE) vehicles by 2025. Not likely, counter Mercedes-Benz's chief technology officer Markus Schäfer and Renault Group CEO Luca de Meo.

The International Energy Agency predicts that EVs will make up more than 60 percent of vehicles sold globally by 2030. But given the sheer tonnage of lithium, cobalt, and other raw materials needed for EV batteries, that figure is overly optimistic, suggests the mineral market analysis company Benchmark Mineral Intelligence, unless nearly 300 new mines and supporting refineries open by then.

EV owners should be urged to charge at night to save not only money and the power grid but "the world," a news headline cries out. Not so fast, exclaim researchers at Stanford University, who state that charging EVs during the day is actually cheaper, better for the grid, and healthier for the environment.

[redacted]

And so goes the litany of contradictory statements about the transition to EVs:

EVs will/will not collapse the electric grid.

EVs will/will not cause massive unemployment among autoworkers.

EVs will/will not create more pollution than they eliminate.

Confused? Join the crowd.

Sorting through this contradictory rhetoric can make anyone's head spin. My response to each proclamation is often a shrug followed by "It depends."

Two years ago, I began investigating the veracity of claims surrounding the transition to EVs at scale. The result is a 12-part series and e-book, *The EV Transition Explained*, that explores the tightly woven technological, policy, and social issues involved. The articles are based on scores of interviews I conducted with managers and engineers in the auto and energy industries, as well as policy experts, academic researchers, market analysts, historians, and EV owners. I also reviewed hundreds of reports, case studies, and books surrounding EVs and electrical grids.

What I found is an intricately tangled web of technological innovation, complexity, and uncertainty, combined with equal amounts of policy optimism and dysfunction. These last two rest on rosy expectations that the public will quietly acquiesce to the considerable disruptions that will inevitably occur in the coming years and decades. The transition to EVs is going to be messier, more expensive, and take far longer than the policymakers who are pushing it believe.

LET ME BE VERY CLEAR: Transitioning to electric vehicles and renewable energy to combat climate change are valid goals in themselves. Drastically reducing our fossil-fuel use is key to realizing those goals. However, attempting to make such transitions at scale in such a short period is fraught with problems, risks, and unanticipated consequences that need honest and open recognition so they can be actively and realistically addressed. Going to scale means not only manufacturing millions of EVs per year but also supporting them from recharging to repair.

A massive effort will be needed to make this happen. For example, in January 2023 the sales of EVs in the United

States reached 7.83 percent of new light-duty-vehicle sales, with 66,416 battery-electric vehicles (BEVs) and 14,143 plug-in hybrid vehicles (PHEVs) sold. But consider that also in January, some 950,000 new ICE light-duty vehicles were sold, as well as approximately another 3 million used ICE vehicles.

Transforming the energy and transportation sectors simultaneously will involve a huge number of known and unknown variables, which will subtly interact in complex, unpredictable ways. As EVs and renewable energy scale up, the problems and the solutions will cover ever-expanding populations and geographies. Each proposed solution will probably create new difficulties. In addition, going to scale threatens people's long-held beliefs, ways of life, and livelihoods, many of which will be altered, if not made obsolete. Technological change is hard, social change even harder.

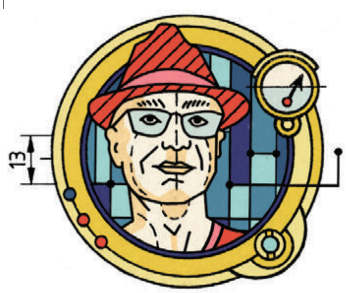
And yet, the rush to transition to EVs is logical. Parts of the world are already experiencing climate-change-related catastrophes, and governments around the world have pledged to act under the Paris Agreement to limit global temperature rise to 1.5 °C above preindustrial levels. This agreement requires the reduction of greenhouse gases across all industrial sectors. Transportation is one of the largest contributors of GHG emissions worldwide, and many experts view replacing ICE vehicles with EVs as being the quickest and easiest way to reach the target of net-zero carbon emissions by 2050.

However, shifting a 125-year-old auto industry that's optimized for ICE-vehicle production to EVs using nascent technology is a monumental challenge in itself. Requiring that automakers do so in 15 years or less is even more daunting, although part of it is their own doing by not recognizing earlier that EVs might be a threat to their business models. EVs require automakers and their suppliers to reinvent their supply chains, hire employees with new software, battery, and mechatronic skill sets, and retrain or else lay off workers whose outdated skills are no longer needed.

The articles in the series address different aspects of this transition, including EV-related unemployment, battery issues, the EV charging infrastructure, and affordability. One not entirely surprising finding is that the

"[GOVERNMENT POLICYMAKERS ASSUME] that they can incentivize the supply and demand of EVs while paying relatively less attention to the capacity of global supply chains to produce them, along with the energy conversion complex needed to power them. Shifting the auto industry, an apex industry supporting a host of others, to meet a new knowledge economy around EVs will be no easy task."

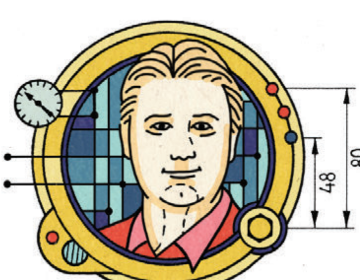
MATTHEW N. EISLER, lecturer at the University of Strathclyde, in Scotland, and a historian and EV expert whose new book is called *Age of Auto Electric*



[redacted]

"WHILE I HAVE often spoken of my concern about the electrical capacity demands required for a significant adoption of EVs, [Charette's] series highlights the many different factors that must change for EVs to be successful. The urgent need to upgrade and increase the number of electric transformers the series discusses, for example, is rarely spoken about."

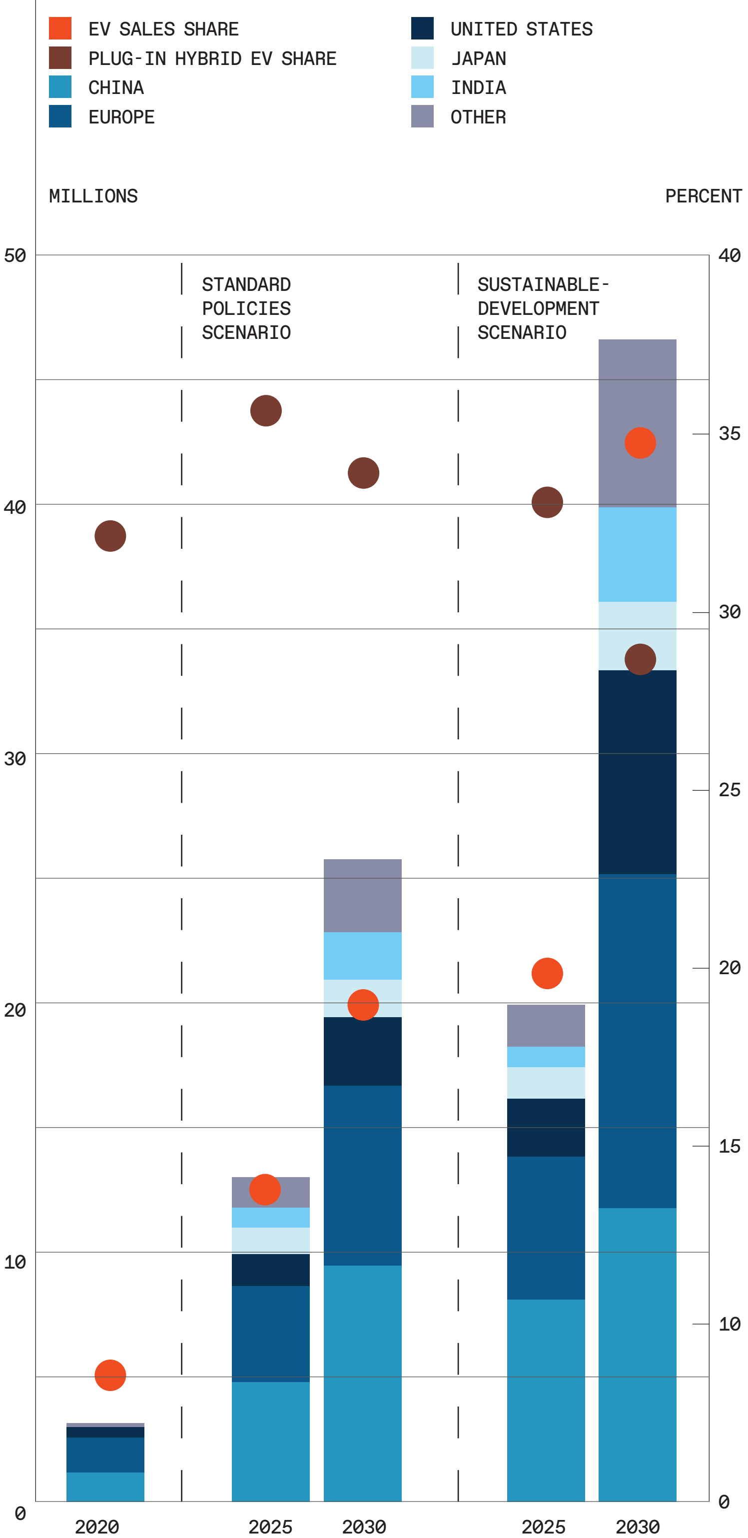
ALEXANDER EDWARDS, president of Strategic Vision, a research consultancy that analyzes how consumers perceive EVs



Portraits by Christian Gralingen

GLOBAL ELECTRIC-VEHICLE SALES BY SCENARIO, 2020-2030

The International Energy Agency's Global EV Outlook 2021 shows 2020 electric-vehicle sales [left column], projected EV sales under current climate-mitigation policies [middle], and projected sales under accelerated climate-mitigation policies [right].



traditional automakers are electrifying their offerings while also squeezing the last bit of profit from their gas guzzlers. That is, they are introducing less-expensive EV models, but their main thrust is still on producing profitable luxury EV models that are well beyond the means of the average household while also pushing sales of profitable fossil-fuel-powered SUVs.

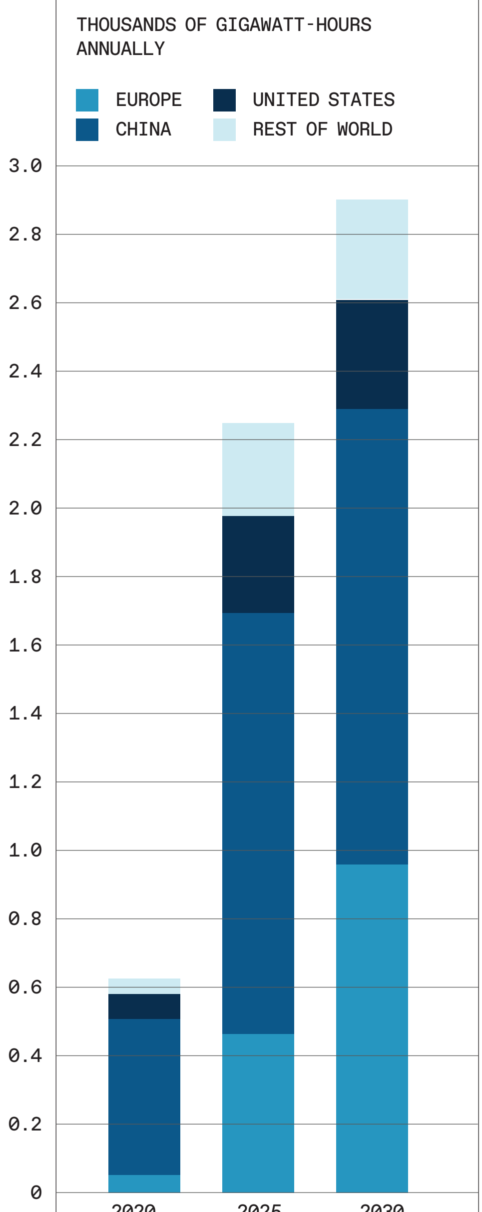
ELECTRIC VEHICLES ARE MORE than just a new technology for combating climate change. In the United States, for instance, policymakers view EVs as the tip of the spear for a vast program of government-directed economic nationalism—the economic, environmental, and societal change aimed at completely reshaping the nation's US $26 trillion economy away from fossil fuels. They see normal market forces as inadequate to meet the imposed climate deadlines. Hence, with the Biden administration's encouragement, ICE-vehicle sales will be banned in 2035 in California and several other states. In the series, I scrutinize several such EV policies and take a look at the roadblocks that could derail them, such as inadequately sized pole transformers and the failure to issue permits for new electricity transmission lines.

The United States is not alone in seeing EVs as an economic driver, of course. Worldwide, nearly 60 countries are now imposing similar ICE-vehicle sales bans. This has forced EVs into yet another role: as a cudgel to be wielded in the fierce geopolitical competition for economic advantage. For China, Japan, the United Kingdom, the European Union, and the United States, EVs are the vehicle needed to "win the future of transportation and manufacturing." Consider the reactions to the recent change in U.S. EV subsidy policy, which aims to boost domestic EV manufacturing and energy security. The decision deeply angered other countries and is sparking moves to counter it.

EVs alone aren't sufficient to meet carbon-reduction targets, which means enormous lifestyle changes for many of us, as we try to do our part to combat climate change. People will need to drive and fly less, walk and bike more, and take public transportation. We'll need to switch to a more plant-based diet and convert household appliances powered by fossil fuels to electricity, to name only

BATTERY-CELL PRODUCTION CAPACITY

This McKinsey & Co. battery tracker from June 2021 shows where batteries are being produced currently and at what volume and where their production is projected to be by the end of the decade. Figures from 2025 and 2030 are estimates based on announcements by battery-cell manufacturers.



SOURCE: MCKINSEY BATTERY TRACKER (JUNE 2021)

Figures from 2025 and 2030 are estimates based on announcements by battery-cell manufacturers.

a few looming adjustments. People's willingness to accept these changes and their ability to implement them will be crucial to our success at adapting to climate change and mitigating its impacts.

The introduction of any new system spawns perturbations that create surprises, both wanted and unwanted. We can safely assume that quickly moving to EVs at scale will unleash its fair share of unpleasant surprises, as well as prove the adage of "haste makes waste."

WHAT STRUCK ME most in writing the series was that the EV transition is incredibly fluid. Major changes in transportation and energy policy, battery technology, and automakers' strategies are announced nearly daily, highlighting the many uncertainties. Given the geopolitical nature of the transition, these uncertainties will only increase.

These rapid changes also show the fragility of the transition. The desperate pleas from automakers for more government subsidies is not reassuring. Tesla's recent price cuts, for instance, have thrown the auto industry into turmoil. Neither is a sign of a market that is sure of itself or its future.

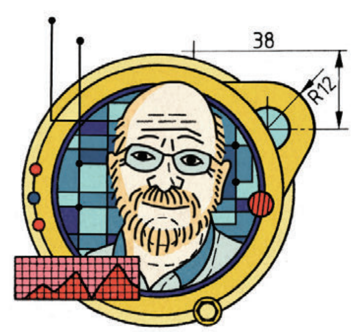
This fragility is also obvious when you examine the overly optimistic assumptions and the many caveats buried in EV and energy-policy recommendations. Many things need to go exactly right, and very little can go wrong for the EV transition to transpire as planned. At times like these, I'm reminded of Nobel Prize–winning physicist Richard Feynman's admonishment: "For a successful technology, reality must take precedence over public relations, for Nature cannot be fooled."

There is a cacophony of foolishness being spouted by those advocating for the EV transition and by those denouncing it. It is time for the nonsense to stop and for some realistic political and systems thinking to begin. ■

Download ***The EV Transition Explained*** e-book, free to IEEE members

"CURRENTLY, THE CADENCE and sequencing of EV policies and engineering activities are out of whack. This creates a concatenation of costly challenges that turns everything you touch into a Pandora's box."

University of Michigan public-policy expert JOHN LESLIE KING



[redacted]

Portrait by Christian Gralingen

False Starts

The checkered history of vehicle-to-grid power

By Matthew N. Eisler

Photo-illustrations by Max-o-matic


Charger

In 2001, a team of engineers at a then-obscure R&D company called AC Propulsion quietly began a groundbreaking experiment. They wanted to see whether an electric vehicle could feed electricity back to the grid. The experiment seemed to prove the feasibility of the technology. The company's president, Tom Gage, dubbed the system "vehicle to grid" or V2G.

The concept behind V2G had gained traction in the late 1990s after California's landmark zero-emission-vehicle (ZEV) mandate went into effect and compelled automakers to commercialize electric cars. In V2G, environmental-policy wonks saw a potent new application of the EV that might satisfy many interests. For the utilities, it promised an economical way of meeting rising demand for electricity. For ratepayers, it offered cheaper and more reliable electricity services. Purveyors of EVs would have a new public-policy rationale backing up their market. And EV owners would become entrepreneurs, selling electricity back to the grid.

AC Propulsion's experiment was timely. It occurred in the wake of the California electricity crisis of 2000 and 2001, when mismanaged deregulation, market manipulation, and environmental catastrophe combined to unhinge the power grid. Some observers thought V2G could prevent the kinds of price spikes and rolling blackouts then plaguing the Golden State. Around the same time, however, General Motors and other automakers were in the process of decommissioning their battery EV fleets, the key component of V2G.

The AC Propulsion experiment thus became an obscure footnote in the tortuous saga of the green automobile. A decade later, in the 2010s, the battery EV began an astounding reversal of fortune, thanks in no small part to the engineers at ACP, whose electric-drive technology informed the development of the Roadster, the car that launched Tesla Motors. By the 2020s, automakers around the world were producing millions of EVs a year. And with the revival of the EV, the V2G concept was reborn.

If a modern electronics-and-software-laden car can be thought of as a computer on wheels, then an electric car capable of discharging electricity to the grid might be considered a sort of power plant on wheels. And indeed, that's how promoters of vehicle-to-grid technology perceive the EV.

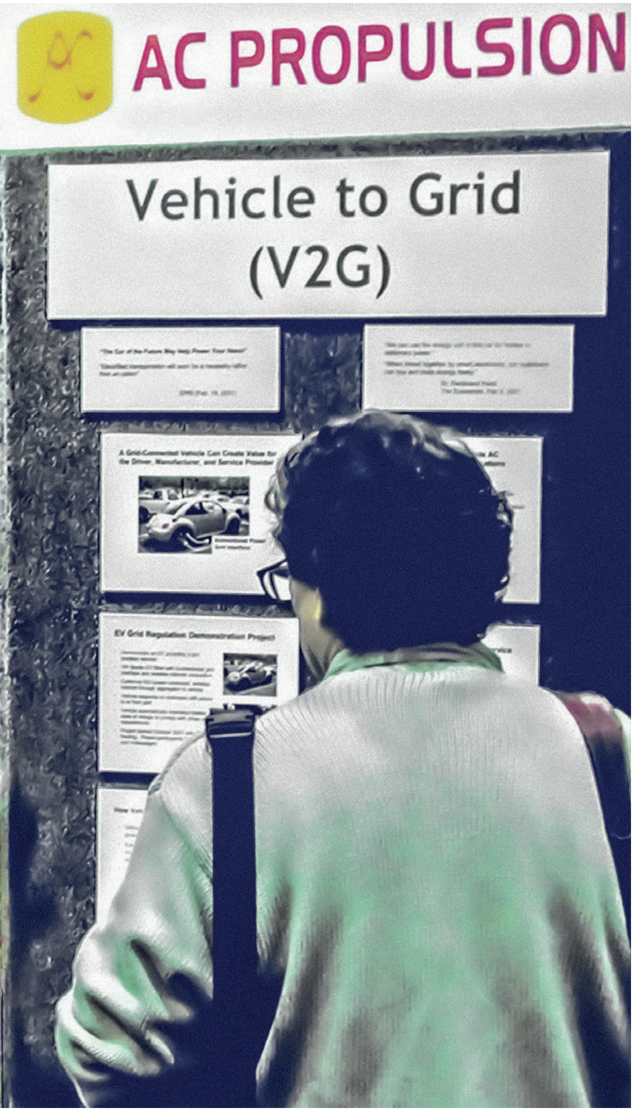
AC Propulsion's president, Tom Gage [above], explains the company's vehicle-to-grid technology at a 2001 conference in Seattle.

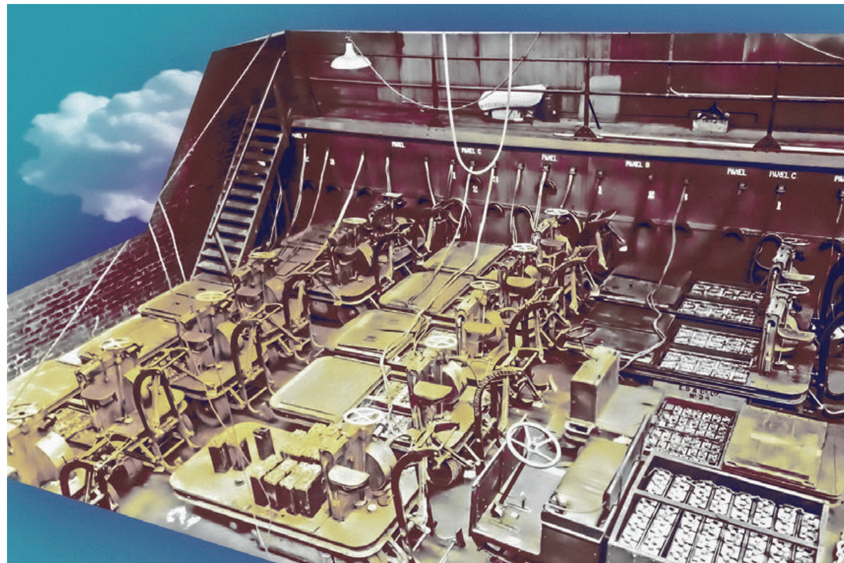
ORIGINAL PHOTOS, ABOVE: ALEC BROOKS (2)

Keep in mind, though, that electricity's unique properties pose problems to anyone who would make a business of producing and delivering it. Electricity is a commodity that is bought and sold, and yet unlike most other commodities, it cannot easily be stored. Once electricity is generated and passes into the grid, it's typically used almost immediately. If too much or too little electricity is present in the power grid, the network can suddenly become unbalanced.

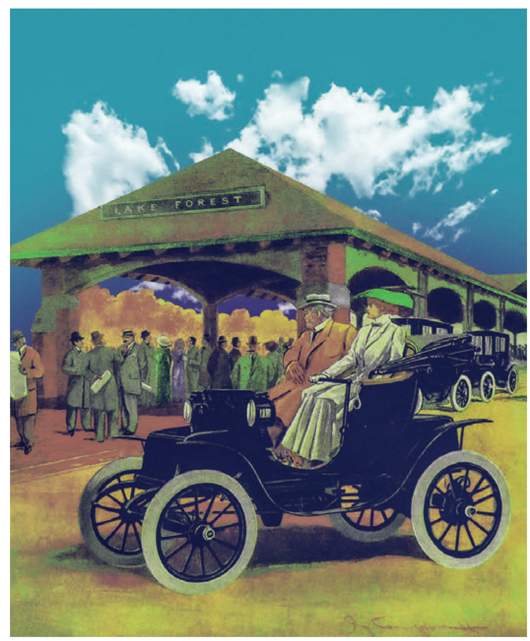
Some operators of early direct-current power plants at the turn of the 20th century solved the problem of uneven power output from their generators by employing large banks of rechargeable lead-acid batteries, which served as a kind of buffer to balance the flow of electrons. As utilities shifted to more reliable alternating-current systems, they phased out these costly backup batteries.

Then, as electricity entrepreneurs expanded power generation and transmission capacity, they





At the turn of the 20th century, utilities promoted the use of electric truck fleets [above] to soak up excess electricity.



In the early years of the automobile, battery-powered electric cars [right] were competitive with cars fueled by gasoline and other types of propulsion.

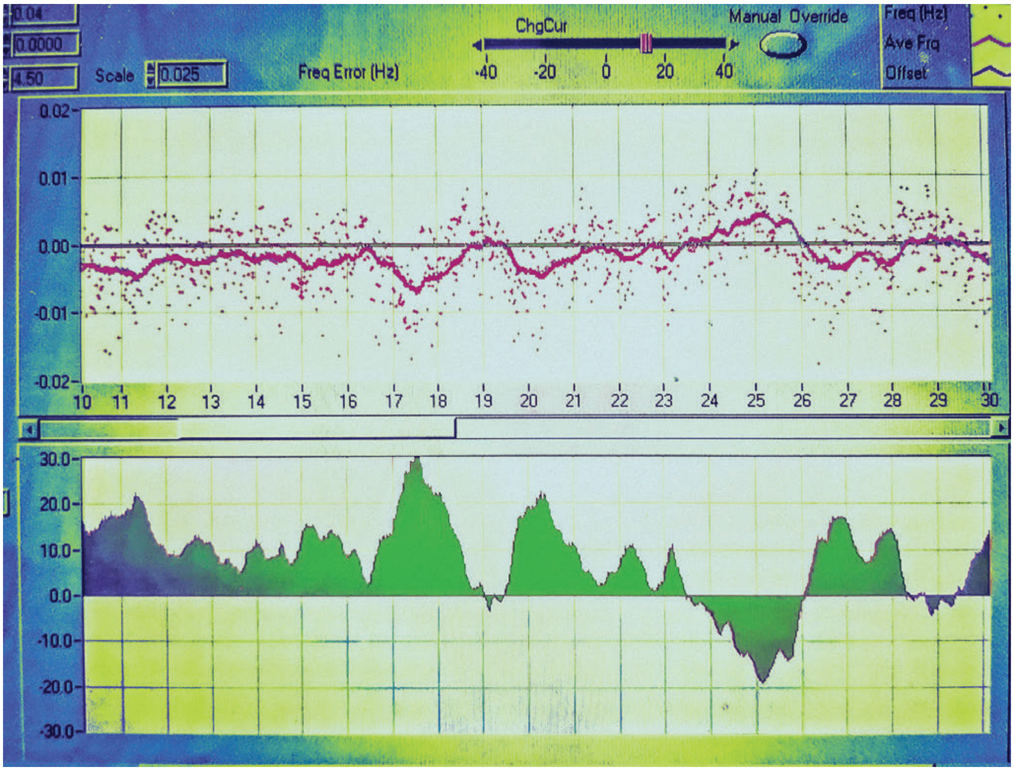
ORIGINAL PHOTOS, TOP RIGHT: M&N/ ALAMY; BOTTOM RIGHT: SHAWSHOTS/ALAMY

faced the new problem of what to do with all the cheap off-peak, nighttime electricity they could now produce. Utilities reconsidered batteries, not as stationary units but in EVs. As the historian Gijs Mom has noted, enterprising utility managers essentially outsourced the storage of electricity to the owners and users of the EVs then proliferating in northeastern U.S. cities. Early utility companies like Boston Edison and New York Edison organized EV fleets, favoring electric trucks for their comparatively capacious batteries.

The problems of grid management that EVs helped solve faded after World War I. In the boom of the 1920s, U.S. utility barons such as Samuel Insull massively expanded the country's grid systems. During the New Deal era, the federal government began funding the construction of giant hydropower plants and pushed transmission into rural areas. By the 1950s, the grid was moving electricity across time zones and national borders, tying in diverse sources of supply and demand.

The need for large-scale electrochemical energy storage as a grid-stabilizing source of demand disappeared. When utilities considered storage technology at all in the succeeding decades, it was generally in the form of pumped-storage hydropower, an expensive piece of infrastructure that could be built only in hilly terrain.

It wasn't until the 1990s that the electric car reemerged as a possible solution to problems of grid electricity. In 1997, Willett Kempton, a professor at the University of Delaware, and Steve Letendre, a professor at Green Mountain College, in Vermont,



began publishing a series of journal articles that imagined the bidirectional EV as a resource for electricity utilities. The researchers estimated that, if applied to the task of generating electricity, all of the engines in the U.S. light-duty-vehicle fleet would produce around 16 times the output of stationary power plants. Kempton and Letendre also noted that the average light vehicle was used only around 4 percent of the time. Therefore, they reasoned, a fleet of bidirectional EVs could be immensely useful to utilities, even if it was only a fraction the size of the conventional vehicle fleet.

The engineers at AC Propulsion (ACP) were familiar with the basic precepts of bidirectional EV power. The company was the brainchild of Wally Rippel and Alan Cocconi, Caltech graduates who had worked in the late 1980s and early 1990s as consultants for AeroVironment, then a developer of lightweight experimental aircraft. The pair made major contributions to the propulsion system for the Impact, a battery-powered concept car that AeroVironment built under contract for General Motors. Forerunner of the famous EV1, the Impact was regarded as the most advanced electric car of its day, thanks to its solid-state power controls, induction motor, and integrated charger. The vehicle inspired California's ZEV mandate, instituted in 1990. As Cocconi told me, the Impact was bidirectional-capable, although that function wasn't fully implemented.

AeroVironment had encouraged its engineers to take creative initiative in developing the Impact,

This pair of graphs shows how the AC-150 drivetrain performed in a demonstration of grid-frequency regulation. The magenta line in the upper graph tracks grid frequency centered around 60 hertz. The lower graph indicates power flowing between the grid and the drivetrain; a negative value means power is being drawn from the grid, while a positive value means power is being sent back to the grid.

but GM tightly managed efforts to translate the idiosyncratic car into a production prototype, which rankled Cocconi and Rippel. Cocconi was also dismayed by the automaker's decision to equip the production car with an off-board rather than onboard charger, which he believed would limit the car's utility. In 1992, he and Rippel quit the project and, with Hughes Aircraft engineer Paul Carosa, founded ACP, to further develop battery-electric propulsion. The team applied their technology to a two-seat sportscar called the tzero, which debuted in January 1997.

Through the 1990s and into the early 2000s, ACP sold its integrated propulsion systems to established automakers, including Honda, Volkswagen, and Volvo, for use in production models being converted into EVs. For car companies, this was a quick and cheap way to gain experience with battery-electric propulsion while also meeting any quota they may have been subject to under the California ZEV mandate.

By the turn of the millennium, however, selling EV propulsion systems had become a hard way to make a living. In early 2000, when GM announced it had ceased production of the EV1, it signaled that the automaking establishment was abandoning battery-electric cars. ACP looked at other ways of marketing

ORIGINAL PHOTOS: LEFT AND ABOVE: ALEC BROOKS

AC Propulsion cofounder Wally Rippel [above] converted a Volkswagen microbus into an electric vehicle while he was still a student at Caltech.

Before cofounding ACP, Alan Cocconi [left] worked on Sunraycer, a solar-powered car for GM. Here, he's testing the car's motor-drive power electronics.

ORIGINAL PHOTO, TOP: HERALD EXAMINER COLLECTION/LAPL

its technology and saw an opportunity in the California electricity crisis then unfolding.

Traditionally, the electricity business combined several discrete services, including some designed to meet demand and others designed to stabilize the network. Since the 1930s, these services had been provided by regulated, vertically integrated utilities, which operated as quasi-monopolies. The most profitable service was peaking power—electricity delivered when demand was highest. The less-lucrative stabilization services balanced electricity load and generation to maintain system frequency at 60 hertz, the standard for the United States. In a vertically integrated utility, peaking services essentially subsidized stabilization services.

With deregulation in the 1990s, these aggregated services were unbundled and commodified. In California, regulators separated generation from distribution and sold 40 percent of installed capacity to newly created independent power producers that specialized in peaking power. Grid-stabilization functions were reborn as "ancillary services." Major utilities were compelled to purchase high-cost peaking power, and because retail prices were capped, they could not pass their costs on to consumers. Moreover, deregulation disincentivized the construction of new power plants. At the turn of the millennium, nearly 20 percent of the state's generating capacity was idled for maintenance.

The newly marketized grid was highly unstable, and in 2000 and 2001, things came to a head. Hot weather caused a demand spike, and the accompanying drought (the beginning of the multidecade southwestern megadrought) cut hydropower capacity. As Californians turned on their air conditioners, peaking capacity had to be kept in operation longer. Then market speculators got into the act, sending wholesale prices up 800 percent and bankrupting Pacific Gas & Electric. Under these combined pressures, grid reliability eroded, resulting in rolling blackouts.

With the grid crippled, ACP's Gage contacted Kempton to discuss whether bidirectional EV power could help. Kempton identified frequency regulation as the optimal V2G market because it was the most profitable of the ancillary services, constituting about 80 percent of what the California Independent System Operator, the nonprofit set up to manage the deregulated grid, then spent on such services.

The result was a demonstration project, a task organized by Alec Brooks, manager of ACP's tzero production. Like Rippel and Cocconi, Brooks was a Caltech graduate and part of the close-knit community of EV enthusiasts that emerged around the prestigious university. After earning a Ph.D. in civil engineering in 1981, Brooks had joined AeroVironment, where he managed the development of Sunraycer, an advanced solar-powered demonstration EV built for GM, and the Impact. He recruited Rippel and Cocconi for both jobs. During the 1990s, Brooks formed a team at AeroVironment that provided support for GM's EV programs until he, too, tired of the corporate routine and joined ACP in 1999.

Working with Gage and Kempton, and consulting with the ISO, Brooks set out to understand how the EV might function as a utility resource.

ACP adapted its second-generation AC-150 drivetrain, which had bidirectional capability, for this application. As Cocconi recalled, the bidirectional function had originally been intended for a different purpose. In the 1990s, batteries had far less capacity than they do today, and for the small community of EV users, the prospect of running out of juice and becoming stranded was very real. In such an emergency, a bidirectional EV with charge to spare could come to the rescue.

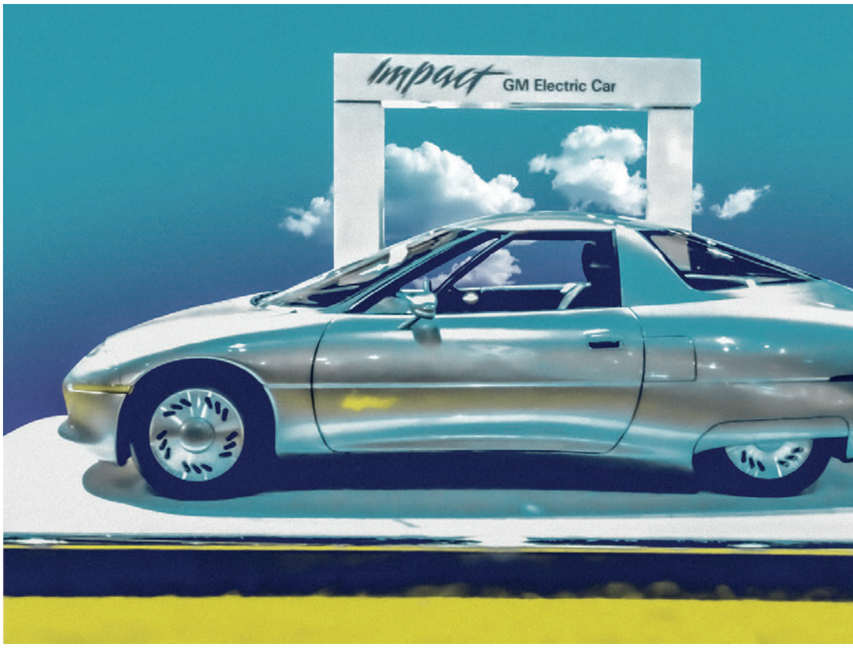
With funding from the California Air Resources Board, the team installed an AC-150 drive in a Volkswagen Beetle. The system converted AC grid power to DC power to charge the battery and could also convert DC power from the battery to AC power that could feed both external stand-alone loads and the grid. Over the course of the project, the group successfully demonstrated bidirectional EV power using simulated dispatch commands from the ISO's computerized energy-management system.

The experiment demonstrated the feasibility of the vehicle-to-grid approach, yet it also revealed the enormous complexities involved in deploying the technology. One unpleasant surprise, Brooks recalled, came with the realization that the electricity crisis had artificially inflated the ancillary-services market. After California resolved the crisis—basically by reregulating and subsidizing electricity—the bubble burst, making frequency regulation using EVs a much less attractive business proposition.

The prospect of integrating EV storage batteries into legacy grid systems also raised concerns about control. The computers responsible for automatically signaling generators to ramp up or down to regulate frequency were programmed to control large thermoelectric and hydroelectric plants, which respond gradually to signals. Batteries, by contrast, respond nearly instantaneously to commands to draw or supply power. David Hawkins, an engineer who served as a chief aide to the ISO's vice president of operations and advised Brooks, noted that the responsiveness of batteries had unintended consequences when they were used to regulate frequency. In one experiment involving a large lithium-ion battery, the control computer fully charged or discharged the unit in a matter of minutes, leaving no spare capacity to regulate the grid.

Electricity's unique properties pose problems to anyone who would make a business of producing and delivering it.

In principle, this problem might have been solved with software to govern the charging and discharging. The main barrier to V2G in the early 2000s, it turns out, was that the battery EV would have to be massively scaled up before it could serve as a practical energy-storage resource. And the auto industry had just canceled the battery EV. In its place, automakers promised the fuel-cell electric car, a type of propulsion system that does not easily lend itself to bidirectional power flow.



General Motors' Impact debuted at the 1990 Los Angeles Auto Show. It was regarded as the most advanced electric vehicle of its era.

The dramatic revival of the battery EV in the late 2000s and early 2010s led by Tesla Motors and Nissan revived prospects for the EV as a power-grid resource. This EV renaissance spawned a host of R&D efforts in bidirectional EV power, including ECOtality and the Mid-Atlantic Grid Interactive Cars Consortium. The consortium, organized by Kempton in conjunction with PJM, the regional transmission organization responsible for much of the eastern United States, used a car equipped with an AC-150 drivetrain to further study the use of V2G in the frequency-regulation market.

Over time, however, the research focus in bidirectional EV applications shifted from the grid to homes and commercial buildings. In the wake of the Fukushima nuclear disaster in 2011, for instance, Nissan developed and marketed a vehicle-to-building (V2B) charging system that enabled its Leaf EV to provide backup power.

The automaker later entered an R&D partnership with Fermata Energy, a Virginia-based company that develops bidirectional EV power systems. Founded by the entrepreneur and University of Virginia researcher David Slutzky in 2010, Fermata considered and then ruled out the frequency-regulation market, on the grounds that it was too small and unscalable.

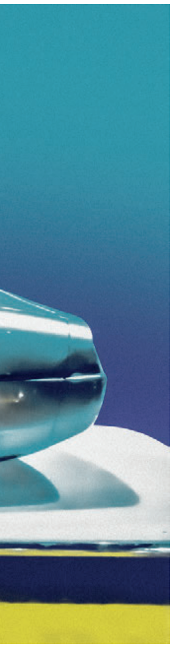
Slutzky now believes that early markets for bidirectional EV power will emerge in supplying backup power and supplementing peak loads for

ORIGINAL PHOTOS: ALEC BROOKS

individual commercial buildings. Those applications will require institutional fleets of EVs. Slutzky and other proponents of EV power have been pressing for a more favorable regulatory environment, including access to the subsidies that states such as California offer to users of stationary storage batteries.

Advocates believe that V2G can help pay for EV batteries. While interest in this idea seems likely to grow as EVs proliferate, the prospect of electric car owners becoming power entrepreneurs appears more distant. Hawkins, the engineer who advised Brooks, holds that the main barriers to V2G are not so much technological as economic: Viable markets need to emerge. The everyday participant in V2G, he argues, would face the difficult task of attempting to arbitrage the difference between wholesale and retail prices while still paying the retail rate. In principle, EV owners could take advantage of the same feed-in tariffs and net-metering schemes designed to enable homeowners to sell surplus solar power back to the grid. But marketing rooftop solar power has proven more complicated and costly for suburbanites than initially assumed, and the same would likely hold true for EV power.

Another major challenge is how to balance the useful lifetime of EV batteries in transportation and nonvehicle applications. That question turns on understanding how EV batteries will perform and age in stationary-power roles. Users would hardly be further ahead, after all, if they substantially degraded their batteries in the act of paying them off. Grid managers could also face problems if they come to depend on EV batteries that prove unreliable or become unavailable as driving patterns change.



In 2001, AC Propulsion engineers installed an AC-150 drivetrain in a Volkswagen Beetle to demonstrate the feasibility of V2G technology for regulating frequency on the power grid.

In short, the core conundrum of V2G is the conflict of interest that comes from repurposing privately owned automobiles as commercial power plants. Scaling up this technology will require intimate collaboration between automaking and electricity-making, enterprises with substantially different revenue models and systems of regulation. At the moment, the auto industry does not have a clear interest in V2G.

On the other hand, rising electricity demand, concerns about fossil fuels, greenhouse gases, and climate change, and the challenges of managing intermittent renewable energy have all created new justifications for bidirectional EV power. With the proliferation of EVs over the last decade, more demonstrations of the technology are being staged for a host of applications—sometimes expressed as V2X, or vehicle-to-everything. Some automakers, notably Nissan and Ford, already sell bidirectional EVs, and others are experimenting with the technology. Enterprises are emerging to equip and manage demonstrations of V2B, V2G, and V2X for utilities and big institutional users of electricity. Some ambitious pilot projects are underway, notably in the Dutch city of Utrecht.

Back in 2002, at the end of their experiment, the engineers at AC Propulsion concluded that what V2G really needed was a powerful institutional champion. They went on to make further important contributions to EV technology. Brooks and Rippel worked for the nascent Tesla Motors, while Cocconi continued at ACP until a cancer diagnosis led him to reevaluate his life. In the mid-2000s, Cocconi sold his stake in the company and devoted himself to aviation, his first love, developing remote-controlled solar-powered aircraft. The rebirth of the battery-electric car in the 2010s and 2020s reaffirmed the efforts of these three visionary pioneers.

A strong V2G patron has yet to emerge. Nevertheless, the idea of an off-the-shelf energy storage unit that also provides transportation and pays for itself is likely to remain attractive enough to sustain ongoing interest. Who knows? The electric car might still one day become a power plant on wheels. ■

The author thanks Alec Brooks, Alan Cocconi, David Hawkins, David Slutzky, and Wally Rippel for sharing their experiences. Parts of this article are adapted from the author's new book, Age of Auto Electric *(MIT Press, 2022).*

BY ALLISON MARSH

Past Forward

This “thunder house” was used to demonstrate the effects of lightning on an unprotected building.

Feel the Thunder

During the 18th and 19th centuries, scientists demonstrated to the public the practicality of Benjamin Franklin’s lightning rod using small model houses—somewhat aptly called thunder houses. The demonstrator would place a small amount of gunpowder inside the house and then zap the house with an electric charge from an early battery called a Leyden jar. If the house had a lightning rod, simulated by a grounded conductor, the charge would pass through without incident. But if the house was set up with an open conductor, the charge would ignite the gunpowder and the sides of the house would collapse with a bang. ■

FOR MORE ON THE HISTORY OF THUNDER HOUSES AND LIGHTNING RODS, SEE spectrum.ieee.org/pastforward-apr2023

NATIONAL MUSEUMS SCOTLAND


YOU'VE INSURED YOUR HOUSE AND CAR.
NOW WHAT ABOUT YOUR LIFE?