STANFORD UNIVERSITY

THE STANFORD EMERGING TECHNOLOGY REVIEW 2026

A Report on Ten Key Technologies and Their Policy Implications

CO-CHAIRS Condoleezza Rice, Jennifer Widom, and Amy Zegart

DIRECTOR AND EDITOR IN CHIEF Herbert S. Lin | **MANAGING EDITOR** Martin Giles

THE STANFORD EMERGING TECHNOLOGY REVIEW 2026

A Report on Ten Key Technologies and Their Policy Implications

CO-CHAIRS

Condoleezza Rice

Jennifer Widom

Amy Zegart

DIRECTOR AND EDITOR IN CHIEF

Herbert S. Lin

MANAGING EDITOR

Martin Giles

Stanford University
Stanford, California

CONTENTS

FOREWORD

This edition of the Stanford *Emerging Technology Review* (*SETR*) coincides with the 250th anniversary of America's Declaration of Independence. As we look toward the future, the past reminds us that history takes surprising turns and that human agency can be powerful. In 1776, few could have dreamed that a ragtag band of colonists in a backwater far from Europe would defeat a great power, replace a king with an extraordinary experiment in democracy, and ultimately become the technological envy of the world. What looked impossible two and a half centuries ago seems inevitable now. Bold ideas and determined action made all the difference.

Today, we face a hinge of history moment where technological discoveries are supercharging both possibility and risk at dizzying speed. This emerging world is hard to understand and even harder to anticipate. But this much seems clear: The choices made today, in everywhere from labs to legislatures, are likely to have consequences for generations. Artificial intelligence (AI) is poised to transform scientific discovery, the future of work, the future of war, and more. And AI is not alone. From nanomaterials that are fifty thousand times smaller than the width of a human hair to commercial satellites and other private-sector technologies deployed in outer space, breakthroughs are reshaping markets, societies, and geopolitics. This is a convergence moment: Never have so many technologies changed so much so fast.

In this era, US technology policy is no longer the unique province of government that it used to be. Federal and state officials are struggling to keep up with technological advances and their implications. At the same time, inventors and investors are struggling to reconcile commercial opportunities and national interests in a world where technology, economics, and geopolitics have become inseparable.

Now more than ever, understanding the landscape of discovery and how to harness technology to forge a better future requires working across sectors, fields, and generations. Engineers and executives need to better understand the policy world to anticipate how their decisions could generate geopolitical advantages and vulnerabilities, and how they can seize opportunities while mitigating risks to the nation. Government leaders need to better understand the academic and business worlds so that well-intended policies don't end up exacerbating societal harms or dampening America's innovation leadership and the geopolitical advantages that come with it. And both government and industry need to better understand the foundational role that America's research universities play in the ecosystem that has made the United States the world's innovation leader since 1945—and how that model is now weakening at home while China is racing to copy it.

The Stanford Emerging Technology Review (SETR) initiative is the first-ever collaboration between Stanford University's School of Engineering, the Hoover Institution, and Stanford's Institute for Human-Centered Artificial Intelligence. We launched this effort with an ambitious goal: transforming technology education for decision makers in both the public and private sectors so that the United States can seize opportunities, mitigate risks, and ensure the American innovation ecosystem continues to thrive.

This is our third annual report surveying the state of ten key emerging technologies and their implications. It harnesses the expertise of leading faculty in science and engineering fields, economics, international relations, and history to identify key technological developments, assess potential implications, and highlight what policymakers should know.

This report is our flagship product, but it is just one element of our continuous technology education campaign for policymakers that now involves more than one hundred Stanford scholars across forty departments and research institutes. In the past year, SETR experts have briefed senior leaders in the private sector and across the US government—in Congress, the White House, the Department of Commerce, the Department of Defense, and the intelligence community. We have organized and participated in dozens of Stanford programs, including multiday congressional staff boot camps in AI, biotechnology, and emerging technologies more broadly; roundtables for CEOs, national media, state and local leaders, and officials from European partners and allies; and workshops convening leaders across sectors to develop new insights that advance space policy, America's biotechnology strategy, defense innovation, and economic statecraft.

Our efforts are guided by three observations:

1. **America's global innovation leadership matters.**

American innovation leadership is not just important for the nation's economy and security. It is the linchpin for maintaining a dynamic global technology innovation ecosystem and securing its benefits for the United States and the world.

Put simply, it matters whether the global innovation ecosystem is led by democracies or autocracies. Democratic countries promote freedom and thrive in it, while authoritarian countries do not. Freedom, in turn, is the fertile soil of innovation, and it takes many forms: the freedom to criticize a government; to admit failure in a research program as a step toward future progress; to share findings openly with others; to collaborate across geographical and technical borders with reciprocal access to talent, knowledge, and resources; and to work without fear of repression, persecution, or political reprisal.

But the United States cannot succeed alone. Robust international collaboration, especially with allies and partners, is essential for bringing together the best minds to tackle the world's toughest challenges, accelerating technological breakthroughs, and advancing American values, not just our interests.

China's rise poses many challenges, and we must not be naive about the Chinese Communist Party's espionage activities and intellectual property theft from American companies and universities or its spread of repressive surveillance technologies around the world. But it is also worth remembering that international scientific collaboration has long been pivotal to fostering global peace, progress, and prosperity, even in times of intense geopolitical competition. During the Cold War, American and Soviet nuclear scientists and policymakers worked together to reduce the risk of accidental nuclear war through arms control agreements and safety measures—at the same time as their nuclear weapons were targeting each other's cities. Similarly, scientific cooperation with China is essential today for reducing shared risks posed by new technologies, from AI-enabled nuclear command and control disasters to conflict in outer space that could bring devastating unintended or unexpected consequences for commercial activities and civilian life.

2. **Academia's role in American innovation is essential—and at risk.**

America's thriving innovation ecosystem has rested on three pillars: the government, the private sector, and the academy. Success has required robust research and development (R&D) in all three. But they are not the same. Evidence suggests that universities' role as the engine of innovation is increasingly at risk, and there is no plan B.

Universities, along with US national laboratories, are the only institutions that conduct research on the frontiers of knowledge without regard for potential profit or foreseeable commercial application. This kind of research is called basic or fundamental research. It takes years, sometimes decades, to bear fruit. And it often fails, because fundamental research is in the business of asking big, hard questions to which nobody knows the answers. But without this kind of research over long periods of time, future commercial innovations would not be possible. Fundamental research investigates questions like, "What are the principles of quantum physics?" and "How does the human immune system work?" Commercial research then builds on openly published academic work to develop quantum computing start-ups whose work could help identify new materials or develop medicines that save millions of lives.

Much of our daily life depends on breakthroughs that would never exist without years of federal investment in fundamental research inside universities. The internet, radar, magnetic resonance imaging (MRI) machines, and the Global Positioning System (GPS) for navigation are just a few examples. Today's AI revolution began fifty years ago with university research into neural networks.

Everyone uses Google, but few people know that Google emerged from a National Science Foundation grant to Stanford professors who were conducting fundamental research on digital libraries back in 1993—when there were one hundred total websites on Earth.[1]

However, there are signs that the engine of innovation in US research universities is not running as well as it could, posing long-term risks to the nation and our technological leadership. In 2024, for the first time, the number of Chinese contributions surpassed ones from the United States in the closely watched *Nature* Index, which tracks eighty-two of the world's premier science journals.[2] Increasingly, the world's best and brightest are not automatically coming to the United States to be educated and possibly stay; global talent has far more educational and training options now than it did ten or twenty years ago. For example, a 2025 Hoover Institution study found that more than half of China's leading AI researchers behind DeepSeek's breakthrough large language model (LLM) were educated and trained entirely in China.[3] In today's technological era, knowledge really is power, and it starts with talent.[4] Reversing the downward slide of American K–12 education at home and recruiting and retaining the brightest minds from abroad have never been more important for American technological competitiveness and national security.

Universities have work to do to fulfill our mission of promoting serious and searching inquiry, restore civic discourse, and regain the trust of the American people. Making cosmetic changes and hoping to return to the way things were will not be enough; this is a moment to reimagine and reinvigorate higher education in service of discovery and the nation. At the same time, the current challenges across US campuses should not distract from the urgent need

Evidence suggests that universities' role as the engine of innovation is increasingly at risk, and there is no plan B.

to ensure American research universities have what it takes to make the breakthrough discoveries of tomorrow. We are harvesting today the research seeds planted decades ago. But we are not planting for the future like we once did.

The US government is the only funder capable of making large and risky investments in the basic science conducted at universities (as well as national laboratories) that is essential for future applications. Yet federal R&D funding has plummeted in percentage terms since the 1960s, from 1.86 percent of GDP in 1964 to just 0.66 percent of GDP in 2016.[5] The United States used to spend more of its GDP on science and research than any nation in the world; today the US ranks eighth.[6]

The Creating Helpful Incentives to Produce Semiconductors (CHIPS) and Science Act of 2022 was supposed to begin reversing this yearslong decline by dramatically increasing federal funding for basic research. But those increases were subsequently scrapped. Current budget proposals call for further reductions in the National Science Foundation budget (which funds all fields of fundamental science and engineering outside of medicine) and the National Institutes of Health budget (which funds medical research).

The United States still funds more basic research than China does, but China is copying the US innovation playbook by investing more and more in basic research and concentrating talent in research universities. In fact, China's basic research investment is rising six times faster than that of the United States. As figure F.1 illustrates, China is poised to overtake the US by the end of the decade if current trends continue.

Private-sector investment in technology companies and associated university research has increased substantially over time, and it may seem like an attractive substitute. But it is not the same. Private investors (rightly) expect returns on their investment, which naturally leads them to fund research avenues with a shorter-term focus and commercial viability. Federal funding for basic research, by contrast, is directed at research that has no foreseeable profit but addresses national issues for the public benefit, seeks to advance basic understanding, and can take a longer-term view to pursue moonshot ideas.[7]

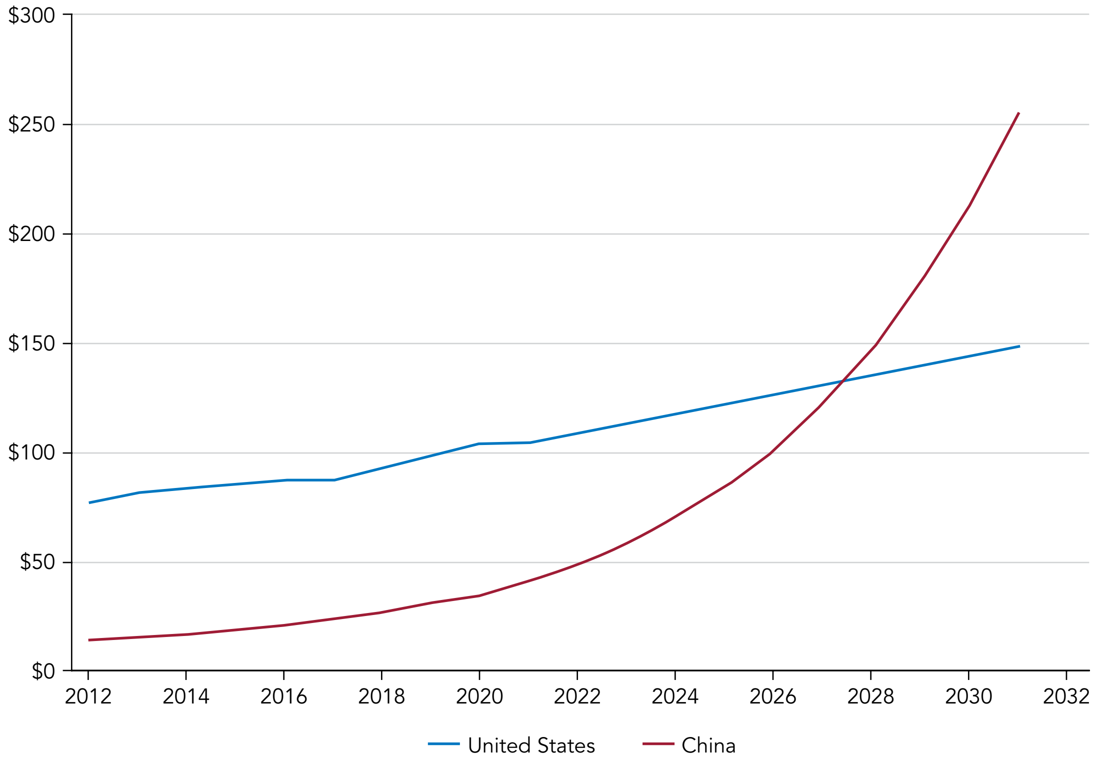
To be sure, the rising dominance of private industry in innovation brings significant benefits. But it is also generating serious and more hidden risks to the health of the entire American innovation ecosystem. In some areas, technology and talent are migrating from academia to the private sector, accelerating the development of commercial products while eroding the foundation for the future. We are already reaching a tipping point in AI. In 2022, more than 70 percent of students who received PhDs in artificial intelligence at US universities took industry jobs, leaving fewer faculty to teach the next generation.[8] As the bipartisan National Security Commission on Artificial Intelligence put it, "Talent follows talent."[9]

Today, only a handful of the world's largest companies have both the talent and the enormous compute power necessary for developing sophisticated LLMs like ChatGPT. No university comes close. In 2024, for example, Princeton University announced that it would tap endowment funds to purchase 300 advanced NVIDIA chips to use for research, costing about $9 million, while Meta announced plans to purchase 350,000 of the same chips by year's end at an estimated cost of $10 billion.[10]

These trends raise several concerning implications.[11] A very significant one is that research in the field is likely to be skewed to applications driven by commercial rather than public interests. The ability for universities—or anyone outside of the leading AI companies—to conduct independent analysis of the weaknesses, risks, and vulnerabilities of AI (especially LLMs recently in the news) will become more important and simultaneously more difficult. Further, the more that industry offers unparalleled talent concentrations, computing power, training data, and the most sophisticated models, the more likely it is that

FIGURE F.1 China is projected to overtake the United States in basic research and development spending

Projected Gross Domestic Expenditure on Basic R&D in Billions of US Dollars



Note: The projection assumes the rate of change between 2012 and 2024 continues forward; it does not include the Trump administration's proposed FY 2026 budget reductions to federally funded research.

Source: OECD Main Science and Technology Indicators Dataset, https://www.oecd.org/en/data/indicators/gross-domestic-spending-on-r-d.html

future generations of the best AI minds will continue to flock there, potentially eroding the nation's ability to conduct broad-ranging foundational research in the field.

3. The view from Stanford is unique, important—and needed now more than ever.

Stanford University has a unique vantage point when it comes to technological innovation. It is not an accident that Silicon Valley surrounds Stanford; technology developed at Stanford in the 1930s served as the foundation for the pioneering companies like Varian Associates and Hewlett-Packard that first shaped industry in the Valley. Since then, the university has continued to fuel that innovation ecosystem. Stanford faculty, researchers, and former students have founded Alphabet, Cisco Systems, Instagram, LinkedIn, NVIDIA, Sun Microsystems, Yahoo!, and many other companies, together generating more annual revenues than most of the world's economies. Start-ups take flight in our dorm rooms, classrooms, kitchens, and laboratories. Technological innovation is lived every day and up close on our campus—with all its benefits and downsides. This ecosystem and its culture, ideas, and perspectives often seem a world apart from the needs and norms

Bridging the divide between the locus of American policy and the heart of American technological innovation has never been more important.

of Washington, DC. Bridging the divide between the locus of American policy and the heart of American technological innovation has never been more important.

Stanford has a rich history of policy engagement, with scholars and alumni who serve at the highest levels of government as well as institutional initiatives that bring together policymakers and researchers to tackle the world's toughest policy problems. And generations of Stanford engineering faculty, students, and staff have had profound impact through their discoveries—from the klystron, a microwave amplifier developed in the 1930s that enabled radar and early satellite communications; to the algorithms driving Google; to optogenetics, a technique pioneered in 2005 that uses light to control neurons, enabling precise studies of brain function. In this moment of technological change, we must do even more to connect emerging technologies with policy. We are proud and excited to continue this unprecedented collaboration to bring policy analysis, social science, science, medicine, and engineering together in new ways.

Today, technology policy and education efforts are often led by policy experts with limited technological expertise. The *Stanford Emerging Technology Review* flips the script, enlisting many of the brightest scientific and engineering minds at the university to share their knowledge of their respective fields by working alongside social scientists to translate their work to nonexpert audiences. We start with science and technology, not policy. And we go from there to emphasize the important interaction between science and all aspects of policy.

How to Use This Report: One Primer, Ten Major Technology Areas

This report is intended to be a one-stop shopping primer that covers developments and implications in ten major emerging technology areas: artificial intelligence; biotechnology and synthetic biology; cryptography and computer security; energy technologies; materials science; neuroscience; quantum technologies; robotics; semiconductors; and space. The list is broad by design, and it includes fields that are widely regarded as pivotal to shaping society, economics, and geopolitics today and into the future.

That said, the ten major technology areas covered in this report are nowhere near an exhaustive catalog of technology research areas at Stanford. And the list may change year to year—not because a particular technology sputtered or we got it wrong, but because categorizing technologies is inherently dynamic; because limiting this report to ten areas imposes discipline on what we cover and how deeply we go; and because we seek to highlight relationships among technologies in ways that may not be obvious. Quantum computing, for example, used to be covered in our chapter on semiconductors, but it is included in a new chapter on quantum technologies this year because of so much current interest in and concern about quantum computing, sensing, and communications. We had a separate chapter on lasers last year, but this year's report folds lasers into our crosscutting themes analysis because the field is more of an enabling technology. Of note, nine of the ten technology chapters appearing in this edition are the same from 2025, and eight of the ten are the same in all three editions of the report.

We have expanded our treatment of issues that cut across technological fields because these are both important and often overlooked. Themes include nonobvious insights that are important for decision makers to remember—like "frontier bias," which is the natural but mistaken assumption that transformational technologies sit on only the frontiers of a field. Indeed, DeepSeek AI's LLM release last year is a cautionary tale that should remind us there are many pathways to success and that not all of them require the most advanced computational resources that American technology firms currently have.

For each of the ten technology chapters, reviews of the field were led by world-renowned, tenured Stanford faculty members who also delivered seminars to faculty contributors, discussants, and SETR advisory board members within and outside their areas of expertise (bios of SETR faculty and contributors can be found at the end of this report). The SETR team also involved more than a dozen postdoctoral scholars and undergraduate research assistants who interviewed faculty across Stanford and drafted background materials.

Each technology chapter begins with an overview of the basics—the major technical subfields, concepts, and terms needed to understand how a technology works and could affect society. Next, we outline important developments and advances in the field. Then we provide an over-the-horizon view of the technology and its future development. Each chapter concludes with a policy section that covers the most crucial considerations for policymakers over the next few years. The report ends with a chapter that looks across the ten technologies, offering analysis of implications for economic growth, national security, environmental and energy sustainability, health and medicine, and civil society.

Three points bear highlighting. First, **we offer no specific policy recommendations in this report**. That is by design. Washington is littered with reports offering policy recommendations that were long forgotten, overtaken by events, or both. Opinions are plentiful. Expert insights based on leading research are not.

We aim to provide a reference resource that is both timeless and timely, an annual state-of-the-art guide that can inform successive generations of policymakers about how to think about evolving technological fields and their implications. Individual SETR faculty may well have views about what should be done. Some of us engage in policy writing and advising. But the mission of this collective report is informing, not advocating. We encourage readers interested in learning more about specific fields and policy ideas to contact our team at SETReview2026@stanford.edu.

Second, **SETR offers *a* view from Stanford, not *the* view from Stanford**. There is no single view of anything in a university. Faculty involved in this report may not agree with everything in it. Their colleagues would probably offer a different lay of the technology landscape with varying assessments about important developments and over-the-horizon issues. This report is intended to reflect an informed judgment about the state of these ten fields—guided by SETR's faculty.

Third, **this report is intended to be the introductory product that translates a broad swath of technological research for nontechnical readers**. Other SETR offerings provide deeper dives into specific technological areas that should be of interest for subject-matter experts.

Ensuring continued American leadership in science and technology is essential, and it's a team effort. We hope this third edition of the *Stanford Emerging Technology Review* continues to spark meaningful dialogue, better policy, and lasting impact. The promise of emerging technology is boundless if, like our founding fathers, we are willing to pursue bold ideas and take determined action.

Condoleezza Rice
Jennifer Widom
Amy Zegart

Co-chairs, Stanford Emerging Technology Review

NOTES

1. "On the Origins of Google," NSF Stories, National Science Foundation, August 17, 2004, https://www.nsf.gov/news/origins-google.

2. Simon Baker, "China Overtakes United States on Contribution to Research in Nature Index," *Nature*, May 19, 2023, https://doi.org/10.1038/d41586-023-01705-7.

3. Amy Zegart and Emerson Johnston, "A Deep Peek into DeepSeek AI's Talent and Implications for US Innovation," Hoover Technology Policy Accelerator / HAI White Paper, April 21, 2025, https://www.hoover.org/research/deep-peek-deepseek-ais-talent-and-implications-us-innovation.

4. Amy Zegart, "The Crumbling Foundations of American Strength," *Foreign Affairs,* September/October 2024.

5. James Manyika and William H. McRaven, *Innovation and National Security: Keeping Our Edge*, Independent Task Force Report No. 77 (Council on Foreign Relations, 2019), 10, https://www.cfr.org/task-force-report/innovation-national-security-keeping-our-edge/cdn/ff/AH68jSeiBTGOynFw4tEuCAb06i8AdHdZ5w9w4QyKLbQ/1676399371/public/2023-02/TFR_Innovation_Strategy.pdf.

6. Abby Joseph Cohen, "The US Risks a Heavy Price for Cutbacks in Research Spending," *Financial Times*, February 25, 2025, https://www.ft.com/content/0d064fab-acf0-4f34-9e3a-a667372f8ef7.

7. Manyika and McRaven, *Innovation and National Security*, 21.

8. Nestor Maslej, Loredana Fattorini, Raymond Perrault, et al., *The AI Index 2024 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI (Stanford University, 2024), 335, https://hai.stanford.edu/ai-index/2024-ai-index-report.

9. Eric Schmidt and Robert Work, *NSCAI Interim Report* (National Security Commission on Artificial Intelligence, 2020), 30, https://apps.dtic.mil/sti/pdfs/AD1112059.pdf.

10. Poornima Apte, "Princeton Invests in New 300-GPU Cluster for Academic AI Research," AI at Princeton, March 15, 2024, https://ai.princeton.edu/news/2024/princeton-invests-new-300-gpu-cluster-academic-ai-research; Michael Kan, "Zuckerberg's Meta Is Spending Billions to Buy 350,000 Nvidia H100 GPUs," *PCMag*, January 18, 2024, https://www.pcmag.com/news/zuckerbergs-meta-is-spending-billions-to-buy-350000-nvidia-h100-gpus; Kif Leswing, "Nvidia's Latest AI Chip Will Cost More than $30,000, CEO Says," CNBC, March 19, 2024, https://www.cnbc.com/2024/03/19/nvidias-blackwell-ai-chip-will-cost-more-than-30000-ceo-says.html.

11. Roman Jurowetzki, Daniel Hain, Juan Mateos-Garcia, and Konstantinos Stathoulopoulos, "The Privatization of AI Research (-ers): Causes and Potential Consequences—From University-Industry Interaction to Public Research Brain-Drain?," preprint, arXiv, 2021, https://arxiv.org/ftp/arxiv/papers/2102/2102.01648.pdf.

EXECUTIVE SUMMARY

Emerging technologies have never been more important or difficult to understand. Breakthrough advances seem to be everywhere, from ChatGPT to the COVID-19 mRNA vaccines to constellations of cheap commercial shoebox-size satellites that can track events on Earth in near-real time. This is a pivotal technological moment offering both tremendous promise and unprecedented challenges. Policymakers need better expert resources to help them understand the burgeoning and complex array of technological developments—more easily and more continuously.

The *Stanford Emerging Technology Review* (*SETR*) is designed to meet this need, offering an easy-to-use reference tool that harnesses the expertise of Stanford University's leading science and engineering faculty in ten major technological areas.

***SETR* 2026 FOCUS TECHNOLOGIES**

Artificial Intelligence
Biotechnology and Synthetic Biology
Cryptography and Computer Security
Energy Technologies
Materials Science
Neuroscience
Quantum Technologies
Robotics
Semiconductors
Space

These particular fields were chosen for this report because they leverage areas of deep expertise at Stanford and cover many critical and emerging technologies identified by the Office of Science and Technology Policy in the White House and by other US government departments. However, *SETR* focus technologies are likely to change over time. This is not because we were incorrect but because science and technology never sleep, the borders between fields are porous, and different people categorize similar research in different ways.

Report Design

This report is organized principally by technology, with each area covered in a stand-alone chapter that gives an overview of the field, highlights key developments, offers an over-the-horizon view of important technological issues, and reviews key policy considerations. Although these chapters can be read individually, one of the most important and unusual hallmarks of this moment is convergence: Emerging technologies are intersecting and interacting in a host of ways, with important implications for policy. We examine these broader dynamics in chapters 11 and 12. In chapter 11, we describe a number of themes and commonalities that cut across many of the technologies we describe earlier in the report. In chapter 12, we consolidate technological developments across all ten areas and discuss how they apply to five policy domains: economic growth, national security, environmental and energy sustainability, health and medicine, and civil society.

Three tensions run throughout and are worth keeping in mind:

- **Timeliness and timelessness** Each chapter seeks to strike a balance between covering recent developments in science and the headlines, and providing essential knowledge about how a field works, what is important within it, and what challenges lie ahead.

- **Technical depth and breadth** This report intentionally skews toward breadth, offering a thirty-thousand-foot view of a vast technological landscape in one compendium. Readers should consider it an introductory course. Other products and educational tools released in the future will offer additional insights into each field.

- **Technical and nontechnical aspects of innovation** We start with the science but do not end with the science. Technological breakthroughs are necessary but not sufficient conditions for successful innovation. Economic, political, and societal factors play enormous and often hidden roles. Johannes Gutenberg invented the printing press in 1452, but it took more than 150 years before the Dutch invented the first successful newspapers. This was not because they perfected the mechanics of movable type but because they decided to use less paper, making newspapers sustainably profitable for the first time. Each chapter in this report was written with an eye toward highlighting important economic, political, policy, legal, and societal factors likely to impede, shape, or accelerate progress.

Technologies and Takeaways at a Glance

Artificial Intelligence

Artificial intelligence (AI) is a computer's ability to perform some of the functions associated with the human brain, including perceiving, reasoning, learning, interacting, problem-solving, and even exercising creativity. In the past year, some of the main AI-related headlines focused on the rapid evolution of the field, including in areas such as large language models and multimodal models integrating vision and language, and on AI's growing adoption by both good and bad actors in society.

KEY CHAPTER TAKEAWAYS

- AI is a foundational technology that is supercharging other scientific fields and, like electricity and the internet, has the potential to transform societies, economies, and politics worldwide.

- Despite rapid progress in the past several years, even the most advanced AI models still have many failure modes and vulnerabilities to cyberattacks that are unpredictable, not widely appreciated nor easily fixed, and capable of leading to unintended consequences.

- Nations are competing to shape the global rules and standards for AI, making interoperability, sizeable national compute resources, and international governance frameworks critical levers of geopolitical influence.

Biotechnology and Synthetic Biology

Biotechnology is the use of cellular and biomolecular processes to develop products or services. Synthetic biology is a subset of biotechnology that involves using engineering tools to modify or create biological functions—like creating a bacterium that can glow in the presence of explosives. Synthetic biology is what created the COVID-19 mRNA vaccine in record time (although the effort relied on decades of earlier research). Just as rockets enabled humans to overcome the constraints of gravity to explore the universe, synthetic biology is enabling humans to overcome the constraints of lineage to develop new living organisms.

KEY CHAPTER TAKEAWAYS

- Biotechnology is emerging as a general-purpose technology by which anything bioengineers learn to encode in DNA can be grown whenever and wherever needed—essentially enabling the production of a wide range of products through biological processes across multiple sectors.
- The United States is still not executing well on strategies for emerging biotechnology and has relied too heavily on private-sector investment to support foundational work needed to scale and sustain progress.
- Biotechnology is one of the most important areas of technological competition between the United States and China, and China is now leveraging two decades of strategic investment to secure global leadership. Absent swift and ambitious actions, the United States risks biotechnological surprise and a loss of biotechnology sovereignty.

Cryptography and Computer Security

The word *cryptography* originates from Greek words that mean "secret writing." In ancient times, cryptography involved the use of ciphers and secret codes. Today, it relies on sophisticated mathematical models to protect data from being altered or accessed inappropriately. Cryptography is often invisible, but it is essential for most internet activities, such as messaging, e-commerce, and banking. In recent years, a type of cryptographic technology called blockchain—which records transactions in distributed ledgers in the computing cloud that cannot be altered retroactively without being detected—has been used for a variety of applications. These include time stamping and ensuring the provenance of information, identity management, supply chain management, and cryptocurrencies.

KEY CHAPTER TAKEAWAYS

- Cryptography is essential for protecting information, but alone it cannot secure cyberspace against all threats; it must operate in concert with the broader field of computer security.
- Cryptography is the enabling technology of blockchain, which is the enabling technology of cryptocurrencies.
- Rather than pursue a central bank digital currency, the United States has adopted a policy preference for privately issued digital assets, promoting stablecoins and cryptocurrencies as vehicles for financial innovation and resilience.

Energy Technologies

Energy is a vital strategic resource for nations that typically involves generation, transmission, and storage. Success in managing energy issues will depend on tackling the "energy trilemma," which is the task of balancing affordability and reliability with reduced greenhouse gas emissions. Energy mix and innovation are key to efforts addressing all three aspects of the trilemma. An important policy issue is achieving greater national consensus about energy goals to enable strategic and effective research and development (R&D) programs and funding.

KEY CHAPTER TAKEAWAYS

- Although many clean energy technologies are now available and increasingly affordable, scaling them up and building the infrastructure for them will take decades due to infrastructure inertia, stakeholder complexity, and the "energy trilemma," which balances reliability, affordability, and cleanliness.

- The US has shifted from climate urgency to energy dominance, redirecting support from renewables and electric vehicles to fission, coal, and natural gas. Globally, similar trends prevail as nations record peak fossil fuel use and scale back renewable investments, prioritizing energy security over decarbonization.

- Energy innovation is fragmented, diverse, and geopolitically strategic, with progress in technologies like fission, geothermal, fusion, and batteries reshaping the energy frontier. To compete with China, US technology leadership depends on sustained R&D funding, robust supply chains, and strategic industrial policies.

Materials Science

Materials science studies the structure and properties of materials—from those visible to the naked eye to microscopic features—and how they can be engineered to change performance. Contributions to the field have led to better semiconductors, "smart bandages" with integrated sensors and simulators to accelerate healing, more easily recyclable plastics, and more energy-efficient solar cells. Materials science has also been key to the development of additive manufacturing, often known as 3-D printing.

KEY CHAPTER TAKEAWAYS

- Materials science is a foundational technology that underlies advances in many other fields, including robotics, space, energy, and synthetic biology.

- The field will exploit artificial intelligence as another promising tool to predict new materials with new properties and to identify novel uses for known materials.

- Future progress in materials science requires new funding mechanisms and access to additional computational power to more effectively transition from innovation to implementation.

Neuroscience

Neuroscience is the study of the human brain and the nervous system—its structure, function, healthy and diseased states, and life cycle from embryonic development to degeneration in later years. The brain is perhaps the least understood and yet most important organ in the human body. Three major research subfields of neuroscience are neuroengineering (e.g., brain–machine interfaces), neurohealth (e.g., brain degeneration and aging), and neurodiscovery (e.g., the science of addiction).

KEY CHAPTER TAKEAWAYS

- Advances in human genetics and experimental neuroscience, along with computing and neuroscience theory, have led to some progress in several areas, including understanding and treating addiction and neurodegenerative diseases, and designing brain–machine interfaces for restoring vision.

- American leadership is essential for establishing and upholding global norms about ethics and human subjects research in neuroscience, but this leadership is slipping with decreased strategic planning and increased foreign investments in the field.

- Popular interest in neuroscience vastly exceeds the current scientific understanding of the brain, giving rise to overhyped claims in the public domain that revolutionary advances are just around the corner.

Quantum Technologies

Quantum technologies exploit the unusual principles of quantum mechanics, such as superposition and entanglement, to create new capabilities in computing, communication, and sensing. Quantum computers are moving toward solving problems that classical systems cannot, with applications in cryptography, materials science, and chemistry. Quantum networking may enable secure communications and scalable computing, while quantum sensors are already advancing navigation, medicine, and environmental monitoring. Though quantum technology (especially computing) is still relatively early in its development, global investment is accelerating, making sustained research and careful policymaking essential to balance innovation, security, and competition.

KEY CHAPTER TAKEAWAYS

- Quantum computing is advancing rapidly, making clear progress toward solving practical problems such as breaking existing public-key encryption algorithms, enabling new materials design, and supporting applications in chemistry. More speculative uses include machine learning, weather modeling, and financial portfolio optimization.

- Quantum networking and sensing are emerging as powerful technologies—networking may be critical for scaling computers to utility levels, while sensors are already transforming fields such as medical imaging and gravitational detection.

- Government-funded basic research in academic labs remains the foundation for breakthroughs, and sustained investment is essential to maintain leadership as companies push applications toward real-world utility.

Robotics

Robotics is an integrative field that draws on advances in multiple technologies rather than a single discipline. The question "What is a robot?" is harder to answer than it appears. At a minimum, the emerging consensus among researchers is that a robot is a physical entity that has ways of sensing itself and the world around it and can create physical effects on that world. Robots are already used across a range of sectors in a variety of ways—including assembly-line manufacturing, space exploration, autonomous vehicles, tele-operated surgery, military reconnaissance, and disaster assistance.

KEY CHAPTER TAKEAWAYS

- Artificial intelligence holds significant potential to advance complex robotic systems, but the speed of future advances will depend on the availability of high-quality training data and the systematic integration of data-rich foundation models, simulated interactions between robots and their environment, and understanding of the real physical world.

- Humanoid robots show promise for specialized industrial and healthcare roles, although widespread adoption of them faces challenges linked to their cost, technical complexity, energy efficiency, safety, and training data quality.

- Advances in autonomous, low-cost, and communication-resilient robotic systems are transforming important aspects of modern warfare.

Semiconductors

Semiconductors, or chips, are crucial and ubiquitous components used in everything from refrigerators and toys to smartphones, cars, computers, and fighter jets. Chip production involves two distinct steps: (1) design, which requires talented engineers to design complex integrated circuits involving millions of components, and (2) fabrication, which is the task of manufacturing chips in large, specially designed factories called "fabs." Because fabs involve highly specialized equipment and facilities, they can cost billions of dollars. US companies still play a leading role in semiconductor design, but capacity for semiconductor manufacturing in America has plummeted, leaving the country heavily dependent on foreign chips, most notably from Taiwan. The Creating Helpful Incentives to Produce Semiconductors (CHIPS) and Science Act of 2022 was intended to help the US semiconductor industry regain a foothold in fabrication, but progress will take years, if not decades.

KEY CHAPTER TAKEAWAYS

- The growing demand for artificial intelligence (AI) and machine learning is driving innovations in chip fabrication, along with advances in memory technologies and high-bandwidth interconnects such as photonic links, all of which are essential for enhancing computational power, managing energy efficiency, and meeting the increasing data needs of modern applications.

- Semiconductor manufacturing is the most precise manufacturing process that exists. It is used to advance work in energy and biotechnology in addition to information technology and AI.

- Strategic technology containment efforts directed against China help constrain Chinese capabilities in the short term. However, they are likely to drive China into a technology posture that is considerably more decoupled from the West and hence less vulnerable to Western pressure in the future.

Space

Space technologies include any technology developed to conduct or support activities approximately sixty miles or more beyond Earth's atmosphere. A single space mission is a system of systems—including everything from the spacecraft itself to propulsion, data storage and processing, electrical power generation and distribution, thermal control to ensure that components are within their operational and survival limits, and ground stations. While in the past, space was the exclusive province of government spy satellites and discovery missions, the number and capabilities of commercial satellites have increased dramatically in recent years. There were roughly one thousand total active satellites in orbit in 2014; today there are around eleven thousand—a figure that will likely rise to several tens of thousands in the next decade.

KEY CHAPTER TAKEAWAYS

- A burgeoning "NewSpace" economy driven by private innovation and investment is transforming space launch, in-space logistics, communications, and key space actors in a domain that until now has been dominated by superpower governments.

- Space is a finite planetary resource. Because of dramatic increases in satellites, debris, and geopolitical space competition, new technologies and new international policy frameworks will be needed to manage the traffic of vehicles, prevent international conflict in space, and ensure responsible stewardship of this global commons.

- The Trump administration has shifted priorities heavily toward human exploration of the Moon and Mars. This is at the expense of robotic exploration, space science, and aeronautics missions, leading to significant planned budget and personnel cuts to NASA. This trend may risk the long-term superiority of the United States in the global race for talent and technology.

Important Crosscutting Themes

Chapter 11 discusses fifteen themes that cut across the technological areas. We split these themes into four categories.

- **Governance and Geopolitics of Emerging Technology** examines how governments and political systems shape global technological progress.
  - Innovation that emerges too fast threatens the legitimate interests of those who might

be negatively affected, while innovation that moves too slowly increases the likelihood that a nation will lose first-mover advantages.

– National monopolies on technology are increasingly difficult to maintain. Even innovations that are solely American born (an increasingly rare occurrence) are unlikely to remain in the exclusive control of American actors for long periods.

– The US government is no longer the primary driver of technological innovation or funder of research and development (R&D).

– While democracies provide greater freedom for scientific exploration, authoritarian regimes can direct sustained funding and focus on technologies they believe are most important.

- **Innovation Pathways and Patterns of Progress** explores the diverse ways in which technological progress unfolds.

  – Technological progress is often unpredictable and nonlinear, with periods of slow development interrupted by sudden breakthroughs. While some fields, like semiconductors, have shown steady improvement, most technologies advance through cycles of experimentation, feedback, and convergence of multiple innovations.

  – Nonscientific factors, such as engineering feasibility, economic viability, manufacturing challenges, and societal acceptance, influence the adoption of technology based on scientific advances.

  – Hype can distort perceptions, leading to inflated expectations that outpace practical utility and distortions in resource allocation.

  – Frontier bias causes overemphasis on new technologies and sometimes results in overlooking impactful uses of established ones.

  – The synergies between different technologies are large and growing, which makes understanding the interactions between different fields all the more important.

- **Human Capital and Knowledge Ecosystems** highlights the critical roles of people, universities, and funding structures in driving and sustaining innovation.

  – Human capital is the foundation of scientific and technological progress. Sustained investment in it is the single most critical factor in ensuring long-term national competitiveness and scientific advancement.

  – Universities are central to both high-risk research and science, technology, engineering, and mathematics (STEM) education. Yet federal R&D funding as a share of GDP has declined, and policy ambiguities hinder international collaboration.

  – The "valley of death" between research feasibility and commercial viability remains a major barrier to advancing innovations to market. New funding models are needed to bridge this gap and sustain America's technological leadership.

- **Infrastructure for Innovation** encompasses vital systems and structures that support innovation on a large scale.

  – Standards enable interoperability, lower costs, and support global trade but can also stifle innovation and be manipulated for market control or geopolitical advantage.

  – Manufacturing is vital for economic resilience and security, especially amid global supply chain disruptions and strategic competition with China and other nations. Technological advances like robotics and artificial intelligence are reshaping production, while policies such as the CHIPS and Science Act of 2022 aim to boost domestic capacity.

– Cybersecurity protects data, systems, and intellectual property from threats, ensuring research integrity and confidentiality. However, maintaining robust security can conflict with the open culture of research environments.

Finally, each of the ten technology fields covered in this report bears on five policy areas that are of interest to policymakers: economic growth, national security, environmental and energy sustainability, health and medicine, and civil society. Chapter 12 identifies applications and consequences of each field as they apply to these policy areas.

INTRODUCTION

The Role of Science and Technology in Advancing National Interests

Vannevar Bush, an engineer and policymaker who oversaw the development of the Manhattan Project, was the nation's first presidential science advisor. In 1945, he wrote, "Advances in science when put to practical use mean more jobs, higher wages, shorter hours, more abundant crops, more leisure for recreation, for study, for learning how to live without the deadening drudgery. . . . Advances in science will also bring higher standards of living, will lead to the prevention or cure of diseases, will promote conservation of our limited national resources, and will assure means of defense against aggression."[1]

Science and technology (S&T) remain essential to our national interests. Advances in them are closely tied to national needs in transportation, agriculture, communication, energy, education, environment, health, and defense—as well as to millions of American jobs. S&T also underpins and drives many strategic objectives in foreign policy, such as reducing the proliferation of weapons of mass destruction, strengthening relationships with allies and partners, improving humanitarian assistance, and promoting growth in developing and transitional economies.[2] Research and development in S&T fields such as information technology, biotechnology, materials science, and space will impact both "hard power" issues—defense, arms control, nonproliferation—and "soft power" concerns, such as climate change, infectious and chronic diseases, and energy supply and demand.[3]

S&T is one important battleground for seeking advantage in geopolitical competition, because advances in scientific and technical fields can contribute to national interests, including a stronger national security posture, greater national pride and self-confidence, economic influence, and diplomatic leverage. But four other points about S&T are equally important:

- Advances in S&T must be accompanied by strong public policy if they are to serve the national interest. Coupling advanced technology with poor policy to influence that technology rarely ends well.
- Advantages gained from S&T advances are transient. Attempting to restrict the transfer of scientific and technical knowledge to other nations may delay its spread, but the first successful demonstration of a technological advance on the part of the United States is often the impetus for other nations to launch their own efforts to catch up.
- Internationally, S&T is not always a zero-sum game, as advances originating in one nation often benefit others. For example, the internet is a US-born innovation whose uses have spread around the world—and the US itself has gained from that spread.
- International competition also occurs with our allies and partners, who engage in the S&T space, developing technology or deploying policy that can leave the United States at a disadvantage.

Policy for S&T

Policymakers have a wide variety of tools to influence the conduct of S&T research and development. Many of these are obvious, such as research funding, tax incentives to firms, intellectual property rights, export controls, classification authority,[4] regulation, public procurement, funding and other aid to strategic sectors, and labor force training and education.

However, policy need not be directed at S&T to meaningfully impact it. For example, immigration policy is not primarily directed at the S&T workforce, but it can have profound effects on the talent available to academic and industry research. Policy oriented in one direction attracts talent to the United States, while policy oriented in another diminishes such talent. Or consider the national economic environment: Stable fiscal and monetary policies make it easier for private-sector decision makers to plan and invest for the long term—a critical consideration when many S&T advances must be nurtured along an extended path from conception to maturity.

Ten S&T Fields

Chapters 1 through 10 describe in more detail ten S&T fields important to the national agenda. Our selection of these fields was driven by several factors: inclusion on common lists of key technologies developed by government, the private sector, and academia, as well as think tanks; and discussions with science and engineering colleagues at Stanford University and other research universities. We do not claim that any one of these ten is more important than any other, and the report addresses them in alphabetical order. Indeed, one of the unexpected aspects of this technological moment is convergence: New technologies are intersecting, overlapping, and driving each other in all sorts of ways—some obvious, some more hidden.

As noted in the foreword, sometimes the technologies covered in the report change. This year's edition of *SETR* includes a new chapter dedicated to quantum technologies. This is substantially longer than the rest of the chapters, but the extra length doesn't mean that this technology area is more important than the others; it simply reflects how the science behind quantum technologies can be hard to grasp and often counterintuitive, even for experts, requiring more words to provide a clear and comprehensible introduction to the subject.

The description of each field is divided into four parts. The first is an **overview** of the field. The second addresses noteworthy **key developments** in the domain that are relevant to understanding the field from a policy perspective. The third, providing an **over-the-horizon** perspective, addresses technological and other developments that are likely to become important in the near future. The last part highlights policy issues that are relevant to the field in question.

In addition to the chapters on individual technologies, this report contains a chapter addressing themes that cut across many—and sometimes all—of the fields in question. These themes came up repeatedly in our conversations with Stanford faculty and provide a broader perspective on how science and society can interact productively for national benefit.

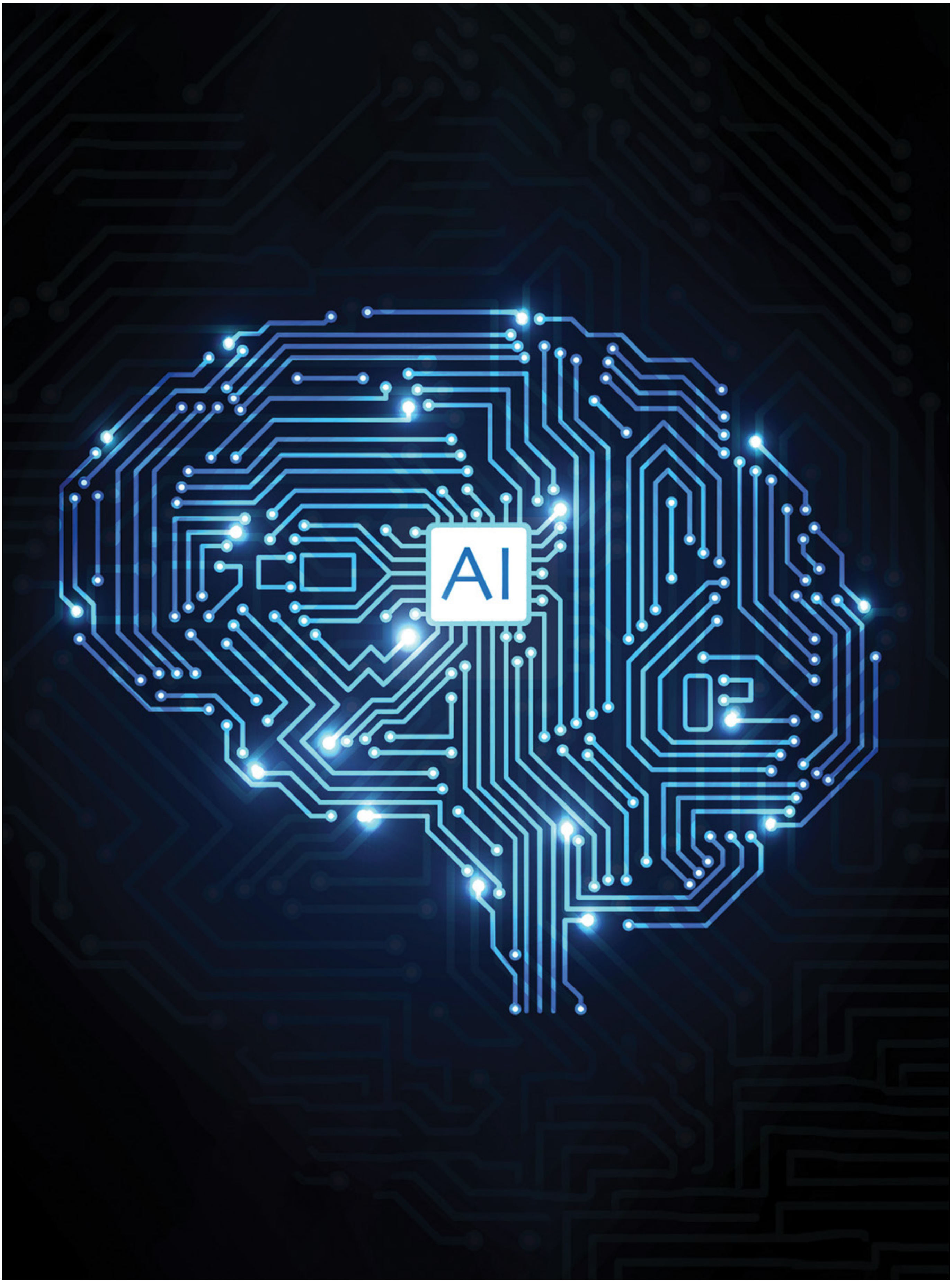
NOTES

1. Vannevar Bush, *Science: The Endless Frontier* (US Government Printing Office, 1945), https://www.nsf.gov/od/lpa/nsf50/vbush1945.htm.

2. National Research Council, *The Pervasive Role of Science, Technology, and Health in Foreign Policy: Imperatives for the Department of State* (National Academies Press, 1999), https://doi.org/10.17226/9688.

3. National Intelligence Council, *Global Trends 2015: A Dialogue About the Future with Nongovernment Officials* (National Intelligence Council, 2000), https://www.dni.gov/files/documents/Global%20Trends_2015%20Report.pdf; National Intelligence Council, *Mapping the Global Future: Report of the National Intelligence Council's 2020 Project* (National Intelligence Council, 2004), https://www.dni.gov/files/documents/Global%20Trends_Mapping%20the%20Global%20Future%202020%20Project.pdf.

4. Under some circumstances (such as when federal funding is involved), the US government may have the authority to classify research even if that research was performed without access to classified information.


AI

01

ARTIFICIAL INTELLIGENCE

KEY TAKEAWAYS

- Artificial intelligence (AI) is a foundational technology that is supercharging other scientific fields and, like electricity and the internet, has the potential to transform societies, economies, and politics worldwide.
- Despite rapid progress in the past several years, even the most advanced AI models still have many failure modes and vulnerabilities to cyberattacks that are unpredictable, not widely appreciated nor easily fixed, and capable of leading to unintended consequences.
- Nations are competing to shape the global rules and standards for AI, making interoperability, sizeable national compute resources, and international governance frameworks critical levers of geopolitical influence.

Overview

Artificial intelligence (AI), a term coined by computer scientist and Stanford professor John McCarthy in 1955, was originally defined as “the science and engineering of making intelligent machines.” In turn, intelligence might be defined as the ability to learn and perform suitable techniques to solve problems and achieve goals appropriate to the context in an uncertain, ever-varying world.[1] AI could be said to refer to a computer’s ability to display this type of intelligence.

The emphasis today in AI is on machines that can learn as well as humans can learn, or at least somewhat comparably so. However, because machines are not limited by the constraints of human biology, AI systems may be able to run at much higher speeds and digest larger volumes and types of information than humans are capable of.

Today, AI promises to be a fundamental enabler of technological advancement in many fields, arguably of comparable importance to electricity in an earlier era or the internet in more recent years. The science of computing, worldwide availability of networks, and civilization-scale data—everything collectively underlying the AI of today and tomorrow—are poised to have similar impact on technological progress in the future. Moreover, the users of AI will not be limited to those with specialized training; instead, the average person on the street will increasingly interact directly with sophisticated AI applications for a multitude of everyday activities.

The global AI market is projected to be worth $244.22 billion in 2025, with North America receiving 33.8 percent of total AI revenues.[2] The Stanford Institute for Human-Centered Artificial Intelligence (HAI) *AI Index 2025 Annual Report* found that private investment in all AI start-ups totaled $150.79 billion in 2024, surpassing the previous record high of over $120 billion, in 2021, after two consecutive years of decline.[3]

One estimate forecasts that generative AI—which can create novel text, images, video, and audio output and is discussed in more detail later in this chapter—could raise global GDP by $7 trillion and raise productivity growth by 1.5 percent over a 10-year period if it is adopted widely.[4] Private funding for generative AI start-ups surged to $33.94 billion in 2024, an 18.7 percent increase from 2023.[5]

The question of what subfields are considered part of AI is a matter of ongoing debate, and the boundaries between these fields are often fluid. Some of the core subfields follow:

- **Machine learning (ML)** Enabling computers to perform tasks without explicit instructions, often by generalizing from patterns in data. This includes deep learning that relies on multilayered artificial neural networks—which process information in a way inspired by the human brain—to model and understand complex relationships within data.
- **Natural language processing** Equipping machines with capabilities to understand, interpret, and produce spoken words and written texts, often using ML techniques.
- **Computer vision** Enabling machines to recognize and understand visual information from the world, convert it into digital data, and make decisions based on these data, often through applying ML.

Much of today's AI is based on ML, though it draws on other subfields as well. ML requires data and computing power—often called compute[6]—and much of today's AI research requires access to these on an enormous scale. The importance of ML is underscored by the award of the 2024 Nobel Prize in Physics (for foundational discoveries enabling ML with artificial neural networks[7]) and the Nobel Prize in Chemistry (for solving the problem of protein structure prediction using AI-based techniques[8]).

In general, a traditional ML model is developed to solve a particular problem, with different problems calling for different models; for problems sufficiently different from each other, entirely new models need to be developed. By contrast, foundation models, a relative newcomer to AI, discussed later in this chapter, may be used across a variety of problems.

ML requires large amounts of data from which it can learn. These data can take various forms, including text, images, videos, sensor readings, and more. Learning from these data is called training the AI model. The quality and quantity of data play a crucial role in determining the performance and capabilities of AI systems. Without sufficient high-quality data, models may generate inaccurate or skewed outcomes. Research continues on how to train systems efficiently: One option is to start from existing models and use a much smaller amount of specially curated data to refine those models' performance for specialized purposes; another option is to compress existing large models into much smaller ones.

For a sense of scale, estimates of the data required to train GPT-4—one of OpenAI's large language models (LLMs), released in March 2023 and the base on which previous versions of ChatGPT were built—suggest that its training database consisted of the textual equivalent of around 100 million books, or about 10 trillion words, drawn from billions of web pages and scanned books. (LLMs are discussed in more detail below.)

The hardware requirements for computing power are also substantial. For example, reports indicate that the training of GPT-4 took about 25,000 Nvidia A100 GPU deep-learning chips—at a cost of $10,000 each—running for about 100 days.[9] Including both these chips and other hardware components used, the overall hardware costs for GPT-4 were at least a few hundred million dollars. And the chips underlying this hardware are specialty ones often fabricated offshore.[10] (Chapter 9, on semiconductors, discusses this point at greater length.)

Lastly, AI models consume a lot of energy. Consider, first, the training phase: One estimate of the electricity required to train a foundation model such as GPT-4 puts the figure at about 50 million kilowatt-hours (kWh).[11] The average American household uses about 11,000 kWh per year, meaning the energy needed to train GPT-4 was approximately the same as that used by 4,500 average homes in a year. Paying for this energy adds significant cost and raises environmental concerns, even before a single person actually uses a model.

Once a model is up and running, the cost of energy used to power queries can add up fast. This is known as the inference phase. For ChatGPT, the energy used per query is around 0.002 of a kilowatt-hour, or 2 watt-hours (Wh).[12] (For comparison, a single Google search requires about 0.3 Wh,[13] and an alkaline AAA battery contains about 2 Wh of energy.)

Given hundreds of millions of queries per day, the operating energy requirement of ChatGPT might be a few hundred thousand kilowatt-hours per day, at a cost of several tens of thousands of dollars. With the recent focus on what are called reasoning models—foundation models that seemingly "think" through problems step by step before presenting the user with an output—such inference costs have substantially increased in the past year.

AI is expected to automate a wide range of tasks. But it also has particular promise in augmenting human capabilities and further enabling people to do what they are best at doing.[14] AI systems can work alongside humans, complementing and assisting their work rather than replacing them. Some present-day examples follow.

Healthcare

- **Medical diagnostics** An AI system that can predict and detect the onset of strokes qualified for Medicare reimbursement in 2020.[15]

- **Drug discovery** An AI-enabled search identified a compound that inhibits the growth of a bacterium responsible for many drug-resistant infections, such as pneumonia and meningitis, by sifting through a library of seven thousand potential drug compounds for an appropriate chemical structure.[16]

- **Patient safety** Smart AI sensors and cameras can improve patient safety in intensive care units, in operating rooms, and even at home by improving healthcare providers' and caregivers' ability to monitor and react to patient health developments, including falls and injuries.[17]

Agriculture

- **Production optimization** AI-enabled computer vision helps some salmon farmers pick out fish that are the right size to keep, thus off-loading the labor-intensive task of sorting them.[18]

- **Crop management** Some farmers are using AI to detect and destroy weeds in a targeted manner, significantly decreasing environmental harm by using herbicides only on undesired vegetation rather than on entire fields. In some cases, this has reduced herbicide use by as much as 90 percent.[19]

Logistics and Transportation

Autonomous trucking From January to August 2024, a consortium of companies autonomously drove trucks carrying tires for over 50,000 long-haul miles.[20] Continued success could automate long-haul drives—the most boring, time-consuming part of a trucker's job—while keeping human tasks like navigating the first miles from factories and the last miles to customers.

Law

Legal review AI-based systems can reduce the time lawyers spend on contract review by as much as 60 percent. Further, such systems can enable lawyers to search case databases more rapidly than online human searches.[21]

Key Developments

Foundation Models

Foundation models have dominated the conversation about AI since late 2022. These models are large-scale systems that are trained on vast amounts of diverse data and that can handle a variety of tasks.[22] They often contain billions or trillions of parameters, and this vast scale allows them to capture more complex patterns and relationships. (Parameters are the building blocks of a foundation model, with values set during training. Though not directly interpretable as discrete chunks of knowledge, parameters act like billions of adjustable knobs that collectively guide how the model learns patterns and makes decisions. The most essential part of a parameter is the weights it uses. Weights represent the strength of connections within the model, encoding learned patterns from training data and determining how much influence each piece of information has on the model's behavior.)

Trained on large-scale data, foundation models can exhibit broad capabilities[23] and are thus sometimes called general-purpose models. They excel at transfer learning—applying knowledge learned in one context to another—making them more flexible and efficient than traditional task-specific models. A single foundation model can be fine-tuned for various tasks, often reducing the need to train separate models from scratch.

These models are generally classified as closed source, open weight, or open source. A closed-source model is a proprietary one developed and maintained by a specific organization, usually a for-profit company, with its source code, data, and architecture kept confidential. Access to these models is typically restricted through technically enforced usage permissions, such as application programming interfaces, allowing the developers to control the model's distribution, usage, and updates.

By contrast, an open-source model is one whose code, data, and underlying architecture are publicly accessible, allowing anyone to use, modify, and distribute it freely. Open-weight models fall in between: Their weights are publicly released, but other components, such as training data, are kept confidential.

The most well-known type of foundation model is an LLM—a system trained on very large volumes of textual content. LLMs are an example of generative AI, a type of AI that can produce new material based on how it has been trained and the inputs it is given. Models trained on text can generate new text based on a statistical analysis that makes predictions about what other words are likely to be found immediately after the occurrence of certain words.

These models do not think or feel like humans do, even though their responses may make it seem like they do. Instead, LLMs use statistical analysis based on training data. For example, because the word sequence "thank you" is far more likely to occur than "thank zebras," a person's query to an LLM asking it to draft a thank-you note to a colleague is unlikely to generate the response "thank zebras."

LLMs generate humanlike language across many subjects, producing potentially useful content such as code, poetry, legal summaries, and medical advice. They outperform median human scores on exams in obstetrics and gynecology;[24] divergent thinking tests;[25] and the LSAT, the GRE, and various Advanced Placement exams.[26] However, they do not necessarily master the underlying skills that these tests assess and still make errors and fail unexpectedly. Developing valid evaluation metrics that accurately capture the true capabilities, limitations, and risks of foundation models remains an open and ongoing research challenge.[27]

Well-known closed-source LLMs include certain of OpenAI's models, such as those in the GPT series; Anthropic's Claude Opus 4.1; and Google's Gemini 2.5 Pro. Well-known open-source or open-weight LLMs include Meta's Llama 4, Google's Gemma 2, and Cohere's Command R.

Specialized foundation models have also been developed in other modalities such as images, audio, and video.

- Foundation models for images are able to generate new images based on a user's text input. Novel methods for handling images, combined with the use of very large collections of pictures and text for training, have led to models that can turn written descriptions into images that are quickly becoming comparable to—and sometimes indistinguishable from—real-life photographs and artwork created by humans. Examples include OpenAI's DALL-E 3, the open-source Stable Diffusion, Google's Imagen, Adobe Firefly, and Meta's Make-A-Scene.
- An example of a foundation model for audio is UniAudio, which handles all audio types and employs predictive algorithms to generate high-quality speech, sound, and music, surpassing leading methods in tasks such as text to speech, speech enhancement, and voice conversion.
- Foundation models in video such as Meta's Emu Video and OpenAI's Sora represent a significant advancement in video generation. Emu first generates an image from a text input and then creates a video based on both the text and the generated image. Sora also enables the user to turn existing images into videos while also editing videos through textual input.

Multimodal Models

AI systems that incorporate multiple modalities—text, images, and sound—within single models are becoming increasingly popular. This multimodal approach, shown in figure 1.1, aims to create more humanlike experiences by leveraging various senses such as sight, speech, and hearing to mirror how humans interact with the world.

Multimodal AI systems have diverse applications across sectors. They can enhance accessibility for people with disabilities through real-time transcription, sign language translation, and detailed image descriptions. They can also eliminate language barriers via cost-effective, near-real-time translation services. In education, multimodal AI can support personalized learning by adapting content to various formats and learner types, improving engagement and comprehension.

When integrated with virtual and augmented reality, AI can create immersive, highly realistic training environments that are particularly valuable in fields like healthcare. The advent of multimodal AI is also set to further transform human–computer interactions, enabling more intuitive communication and expanding the range of tasks that AI systems can handle.

Limitations and Risks of Current AI Systems

Potential positive impacts of AI will likely come from societal applications. But no technology is an unalloyed good. Negative impacts are expected both from current AI limitations and future advances. Key issues include the following:

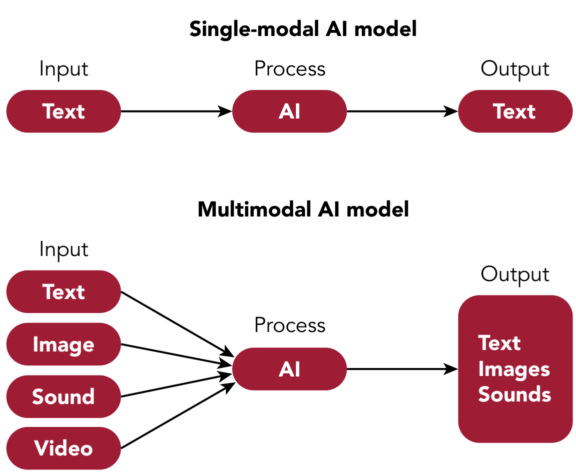
- **Explainability** AI systems generally cannot explain their reasoning or data sources. While explanations aren't always needed, in critical domains, like medical decision making, they are essential for user confidence and trust.

- **Bias and fairness** Models trained on biased datasets reproduce those biases. For example, a facial recognition system trained mainly on one ethnic group may perform poorly on others, likely leading to disproportionate harms.[28] Because data reflects historical inequities, models inevitably embed them, too.

- **Vulnerability to attacks** Small changes to data or inputs can trick AI into false conclusions. For example, small changes, invisible to the naked eye, made to the individual pixels of a stop sign image can cause an AI to classify it as a yield sign.[29] (Chapter 3, on cryptography and computer security, discusses this point in greater detail.) This could prove particularly dangerous for systems used in medicine or the military. Newer models (e.g., multimodal models and agents, covered later in this chapter) expand possible attack vectors.

- **Deepfakes** AI provides the capability for generating highly realistic but entirely inauthentic audio and video imagery. This has obvious implications for evidence presented in courtrooms and for efforts to manipulate political contests. However, despite widespread concerns expressed in 2024 about the potential effects of deepfakes on elections, fake audio, images, and videos did not play as transformative or disruptive a role as feared in the 2024 US elections.[30] Of particular interest is the observation that traditional "cheap fakes"—relatively crude attempts made by manipulating videos and other content—were more prevalent than AI-generated deepfakes. Nevertheless, concerns remain about the potential impact on future democratic processes as the sophistication and usage of AI-generated deepfakes increase.

FIGURE 1.1 Multimodal AI systems can transform one type of input into a different type of output



AI agents ideally operate by executing tasks with minimal human input and oversight. . . . Yet, from a technical standpoint, present-day [agents] face major limitations.

- **Privacy** Many LLMs are trained on vast amounts of internet data, often without careful filtering, and such data can include individuals' personal information. Once incorporated into training datasets, this information may be reproduced or disclosed by the model. Additionally, as AI handles sensitive tasks like mental health support, privacy concerns will grow.

- **Overtrust and overreliance** Familiarity increases user trust, but people may become too complacent when employing AI tools. For example, a recent study showed that developers who used AI coding assistants wrote less secure code—yet they believed that what they were producing was more secure.[31]

- **Hallucinations** Hallucinations occur whenever models generate plausible but false outputs, leaving users unaware that the outputs are fabricated. In September 2024, a Stanford professor asked an AI to list ten of her publications. It returned five real publications and five invented ones, complete with convincing titles and summaries. When she flagged the errors as hallucinations, the model simply produced two new fabricated results.

Researchers are aware of these problems and are working on fixes, but solutions often don't generalize beyond specific cases.

Over the Horizon

AI Agents

AI agents ideally operate by executing tasks with minimal human input and oversight, such as setting people's daily agendas and coordinating software tools. Such agents are gaining traction in activities within enterprises such as customer service and invoice processing. OpenAI's strategy for ChatGPT involves it becoming a "super assistant" that handles everyday tasks such as drafting emails, offering medical advice, and managing finances.[32] Yet, from a technical standpoint, present-day AI agents face major limitations.

- **Memory** The basic unit that an LLM reads and generates to process text—whether a whole word, part of a word, or punctuation—is known as a token. For AI agents to be effective, they need to remember things, such as when a meeting was scheduled in a previous session. An agent's memory is limited by context length, which is the maximum number of input and output tokens the system can handle at any one time. Although the context length of top systems has expanded dramatically in recent years, it is still not enough to remember all the details needed to execute many multistep tasks, especially across different sessions. Efforts to increase cross-session memory and enhance long-term storage are still nascent.

- **Reliability** Even with adequate memory, agents can suffer from goal drift, infinite loops, and resource exhaustion, which undermine their usefulness in real-world settings.
  - Goal drift occurs when an AI agent fails to consistently pursue its original objective during task execution and instead begins to focus on less relevant objectives, which can result in undesired behavior or outcomes.
  - Infinite loops occur when an agent "gets stuck," repeatedly performing the same actions or reasoning steps without making progress toward its goal.
  - Resource exhaustion happens when an agent consumes excessive computational or memory resources due to complex processing, repeated retries, or ineffective algorithms, resulting in degraded performance or system failure.
- **Interoperability** Most agents cannot seamlessly communicate with other agents or external systems. However, in November 2024, the Model Context Protocol (MCP), an open standard for secure, efficient agent-to-system integration, was introduced by Anthropic; since then, it has been adopted by OpenAI, Google DeepMind, Microsoft, and others. The MCP provides a universal interface for an agent to read files, execute functions, handle contextual prompts, and connect with external tools, data sources, and applications, enabling AI agents to access real-time data and perform actions beyond their training data.

Efficiency, Specialization, and Synthetic Data

Progress in AI is shifting from building ever more resource-intensive models to using resources more efficiently. For example, because of limited supplies of real data, synthetic data is increasingly used. (In this context, synthetic data is artificially generated data that are designed to mimic the statistical properties and patterns of real-world data. However, such data do not themselves reveal anything real about the world.)

In addition, because large models consume significant amounts of energy, AI engineers are focusing on models that can be constructed with fewer computational resources. For example, one approach is to do individual calculations at a lower level of precision (that is, using fewer bits to represent a number), thus reducing the volume of data processed and memory used. Individual calculations are therefore less accurate, but the model is trained using a huge number of calculations on different though similar data; averaging over many such calculations can compensate for the loss of accuracy in individual ones. A variety of other approaches also rely on eliminating parts of model training that may be unnecessary or redundant for a given purpose.

Future AI gains will increasingly depend not just on large compute capacity and large amounts of data but also on domain-specific data and efficiency-focused innovations. Quantum computing (discussed in chapter 7, on quantum technologies) may lower the requirements of ML for energy and compute, although it is unclear whether such enhancements are within reach of current technologies.[33]

Embodied AI

Embodied AI means AI that is able to sense and act in the real world (e.g., through its integration into robots or other physical devices). This has the potential to enhance robotic capabilities and expand the range of interactions robots have with the physical world. Systems combining robots and AI could potentially address knowledge tasks, physical tasks, or combinations of both. (This topic is explored further in chapter 8, on robotics.) As research progresses in AI autonomy and reasoning, embodied AI systems may be able to handle increasingly complex tasks with greater independence. This could lead to applications in various fields such as logistics and domestic assistance.

Policy Issues

The Future of Work

Within five to ten years, more workers will have AI integrated into their workflows or will have their jobs replaced entirely by AI systems, a potential disruption to the job market.[34] LLMs have already shown they are sometimes useful in fields like law, customer support, coding, and journalism. These developments raise concerns about AI's significant impact on many knowledge-based jobs and employment overall. However, uncertainty abounds. What and how many present-day jobs will disappear? Which tasks could best be handled by AI? And what new jobs might be created by the technology today and in the future?

Some broad outlines and trends are clear.

- Individuals whose jobs entail routine white-collar work may be more affected than those whose jobs require physical labor; some will experience painful shifts in the short term.[35]

- AI is helping some workers to increase their productivity and job satisfaction.[36] At the same time, other workers are already losing their jobs as AI demonstrates some competence for business operations—despite potentially underperforming the humans it replaces.[37] In some cases, companies are deciding that the cost savings of eliminating human workers outweigh the drawbacks of mediocre AI performance.

- Training displaced workers to be more competitive in an AI-enabled economy does not solve the problem if new jobs are not available. The nature and extent of new roles resulting from widespread AI deployment are not clear at this point. However, historically, the introduction of new technologies has not resulted in a long-term net loss of jobs.[38]

Governance and Regulation of AI

Governments around the world have been increasingly focused on establishing regulations and guidelines for AI.

In the United States, the Trump administration revoked the Biden administration's Executive Order (EO) 14110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. It subsequently issued a new EO titled Removing Barriers to American Leadership in Artificial Intelligence, aiming to promote AI innovation and leadership by eliminating the restrictions and requirements on AI contained in EO 14110. In August 2025, the administration launched America's AI Action Plan, which outlines a policy road map to "accelerate innovation, build American AI infrastructure, and lead in international diplomacy and security."[39] This plan faces challenges, especially regarding alignment with concurrent proposals to scale back broader scientific research funding.

In addition, friction between state and federal approaches to AI governance is growing. States are experimenting with their own AI legislation, often proposing requirements that go well beyond federal guidance. Advocates view this "policy laboratory" approach as essential for innovation in governance; critics warn it creates a fragmented compliance landscape that hampers interstate commerce. Recent state action, as described in table 1.1, underscores the emerging rule-making patchwork.

In the European Union, the most ambitious attempts to regulate AI came into force in August 2024 with the EU Artificial Intelligence Act. The act forbids certain applications of AI, such as individual predictive policing based solely on a person's data profile or tracking of their emotional state in the workplace and educational institutions, unless for medical or safety reasons.[40] Additionally, it imposes a number of requirements on high-risk AI systems and foundation models, addressing transparency and explainability, human oversight, cybersecurity, and robustness.

TABLE 1.1 Selected states' actions regarding AI legislation

State	Legislation	Key provisions
Colorado	Senate Bill 24-205 (Colorado AI Act)	Mandates duties on developers and deployers of "high-risk" AI systems to prevent algorithmic discrimination
Texas	Texas Responsible Artificial Intelligence Governance Act	Prohibits AI systems used for behavioral manipulation, discrimination, or deployment of deepfakes
California	Multiple AI bills (15+), including Assembly Bill 2013	Targets generative AI systems and requires training data disclosure for AI systems used by Californians

To help operationalize these rules, the European AI Office oversaw one of the earliest and most formalized multi-stakeholder consultations in AI policy to date: the General-Purpose AI (GPAI) Code of Practice. Developed by thirteen independent experts with input from nearly one thousand participants across member states, academia, civil society, and industry, the code supplements the EU AI Act by offering detailed provisions on transparency, copyright, and safety and security—giving foundation model developers a recognized pathway to meeting certain requirements of the EU AI Act with greater legal clarity and reduced enforcement risk.

Other important international AI governance developments include the AI Summit series. The first AI Safety Summit, held in November 2023 at Bletchley Park in the United Kingdom,[41] issued the Bletchley Declaration. In it, the European Union and twenty-eight nations collectively endorsed international cooperation to manage risks associated with highly capable general-purpose AI models. The summit also led to the establishment of the United Kingdom's AI Safety Institute (now the UK AI Security Institute) and the US AI Safety Institute (now the Center for AI Standards and Innovation), located within the National Institute of Standards and Technology. Similar institutions have since been established in a number of other countries, including Japan, Singapore, South Korea, Canada, France, Kenya, Australia, and the European Union.

The Seoul Declaration, from the AI Seoul Summit 2024, built on the Bletchley Declaration to acknowledge the importance of interoperability between national AI governance frameworks to maximize benefits and minimize risks from advanced AI systems. In contrast, the 2025 Artificial Intelligence Action Summit, in France, marked a notable shift from the safety-oriented tone of earlier gatherings, placing greater emphasis on accelerating innovation and industrial adoption of AI.

National Security and Geopolitics

The intensifying technological race between the United States and China regarding AI is entering a new phase. While the United States continues to push the technical frontier with increasingly capable models, China is aggressively diffusing existing AI capabilities across every sector—from education to manufacturing to governance—aiming to lock in large-scale network advantages at home and abroad. China's open-source model releases, such as DeepSeek, further challenge America's frontier status, accelerate global adoption, and undermine US containment efforts.

Infrastructure—especially compute capacity—has become a critical, strategic global resource that is vital to remaining competitive in economic, military, cyber, and intelligence domains. For example, countries like Canada and the United Kingdom have announced major compute infrastructure projects.[42] In the United States, the privately funded Stargate AI infrastructure initiative was launched in January 2025. First reports suggested this would invest as much as $500 billion over the next few years, although a more recent report suggests a scaling back of initial objectives.[43]

The United States also began piloting the National Artificial Intelligence Research Resource (NAIRR) in January 2024. Supported by the Trump administration, NAIRR is a federally backed initiative to give academics and civil-society researchers shared access to advanced compute, high-quality data, and AI tools. However, it is noteworthy that investments for AI from high-tech companies exceeded $27 billion in 2023 alone. This figure is far larger than the $2.6 billion authorized for appropriation over six years for NAIRR under the Creating Resources for Every American To Experiment with Artificial Intelligence (CREATE AI) Act of 2025.

Export Controls

Export controls are often used to slow the advances of rivals, but open-source diffusion and breakthroughs in compute-efficient training may erode their impact. This leaves policymakers with a strategic choice: focus on restraining competitors through export controls, or accelerate domestic innovation and the global adoption of domestic products—or attempt both simultaneously. Under the Biden administration, and initially under the Trump administration, the United States has taken the export control route. However, in August 2025 reports suggested a shift, with President Trump and Chinese President Xi Jinping considering an arrangement allowing US chipmakers Nvidia and AMD (Advanced Micro Devices) to sell certain chips to China in exchange for a 15 percent revenue share to the US government.[44]

Use in the Military

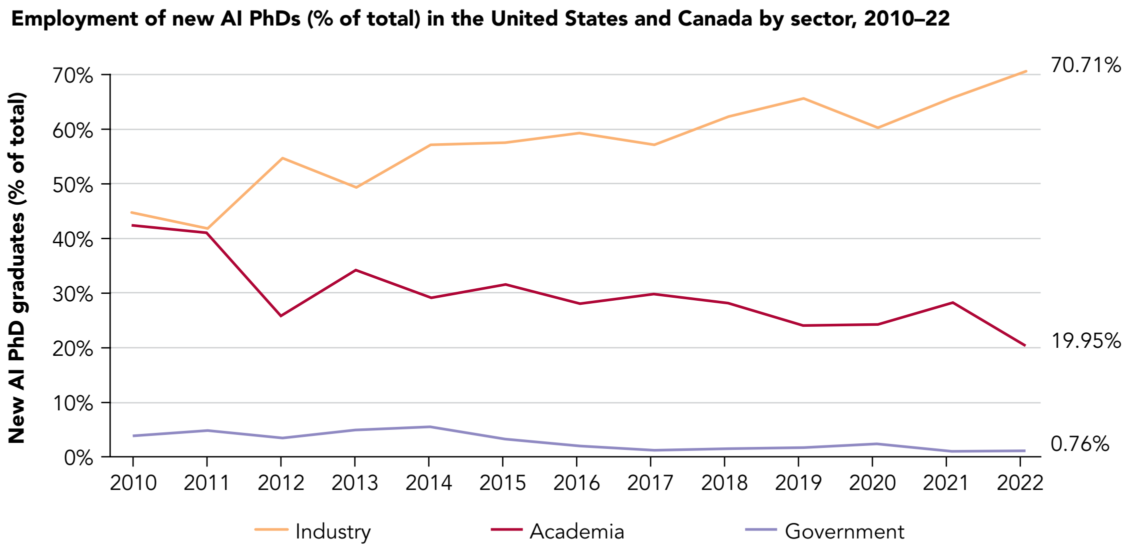
AI is expected to have a profound impact on militaries worldwide.[45] Weapons systems, command and control, logistics, acquisition, and training will all seek to leverage AI to operate more effectively and efficiently, at lower cost and with less risk to friendly forces. Trying to overcome decades of institutional inertia, the US Department of Defense is dedicating billions of dollars to institutional reforms and research advances aimed at integrating AI into its warfighting and war preparation strategies. Senior military officials are concerned that failure to adapt to the emerging opportunities and challenges presented by AI would pose significant national security risks, particularly considering that both Russia and China are investing heavily in AI capabilities.

Talent

The United States is eating its seed corn with respect to the AI talent pool. Faculty at Stanford and other universities report that the number of students studying in AI who are joining industry, particularly startups, is increasing at the expense of those pursuing academic careers and contributing to foundational AI

The resources needed to train GPT-4 far exceed those available through grants or any other sources to any reasonably sized group of the top US research universities.

FIGURE 1.2 Most new AI PhDs hired in North America are flocking to industry



Source: Adapted from Nestor Maslej, Loredana Fattorini, Raymond Perrault, et al., *The AI Index 2024 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024. Data from CRA Taulbee Survey, 2023

research. The United States is thus experiencing an AI "brain drain" that does not favor the US research enterprise or its innovation capacity (figure 1.2).

Many factors are contributing to this trend. One is that industry careers offer compensation far exceeding academic packages. Academic researchers must also secure money for equipment, compute, and staff, often relying on government funding that is typically small compared to what large companies might be willing to invest in their own researchers. Consider, for example, that the resources needed to build and train GPT-4 far exceed those available through grants or any other sources to any reasonably sized group of the top US research universities, let alone any single university. This gap is exacerbated by recent cuts in federal research funding.

Industry often makes decisions more rapidly than government grant makers and imposes fewer regulations on the conduct of research. Large companies are also at an advantage because they have research-supporting infrastructure in place, such as compute facilities and data warehouses.

Finally, other nations are actively recruiting talent that in the past tended to favor employment in the United States. China's recruitment of top scientific talent, especially ethnic Chinese in the United States, is being driven through offers of benefits such as high salaries and generous research funding. Countries like Canada and the United Kingdom actively recruit US-based researchers and offer AI-focused visas. The brain drain has been exacerbated by changes in US immigration policy that have caused top researchers to leave and deterred talented international students from studying in the United States.

Copyright

Many foundation models have been trained on vast amounts of data found on the internet. These data have generally been used without the consent or permission of their owners, and a number of lawsuits were filed over this issue in 2023–24. They include Getty Images v. Stability AI over alleged infringement on copyrights of photographs;[46] *The New York Times* v. OpenAI and Microsoft over alleged use of millions of articles published by the *Times*;[47] and Sony Music, Universal Music Group, and Warner Records v. AI start-ups Suno and Udio over alleged use of protected content to train their music-generation systems.[48] Such cases have raised important questions about appropriately compensating and acknowledging data creators whose data is used to train AI models.

At the time of this writing (September 2025), one major copyright lawsuit against an AI company has reached a settlement. Anthropic, an AI developer, agreed to pay $1.5 billion to settle a class-action suit brought by authors and publishers;[49] the settlement was approved on a preliminary basis by the judge overseeing the case as this publication went to press.[50] Even if it is approved, it will not create a legal precedent because of the out-of-court nature of the settlement. Whether the use of publicly available web data to train AI models is legally permissible remains, and will remain, unsettled across jurisdictions until more court decisions are rendered.

Glossary

Compute: The processing power, typically measured in number of specialized chips and scale of energy use, required to train and run AI models.

Distillation: A technique for compressing a large, complex AI model into a smaller one that is faster and more efficient while retaining most of its capabilities.

Generative AI / Foundation models: Large, general-purpose AI systems trained on vast datasets that can generate text, images, code, or other outputs and be adapted to many downstream applications.

Inference: The stage when a trained AI model is used to generate predictions, outputs, or decisions in response to new inputs.

Model training: The process of teaching an AI system by exposing it to large datasets and adjusting its parameters until it can perform a given task (or a set of tasks) effectively.

Multimodal AI: AI systems designed to process and integrate multiple kinds of data such as text, images, audio, or video into a single model.

Open source / Open weight / Closed source: Terms describing how freely AI models are shared; open-source makes training data, code, and weights public; open-weight shares only the trained parameters; and closed-source restricts access entirely.

Scaling laws: Observed patterns showing that, as AI models are trained with more data, parameters, and compute, their performance improves in predictable ways but also with potentially diminishing returns and real-world limits.

Synthetic data: Artificially generated data such as simulated text, images, or code used to supplement or replace real-world data for training and testing AI systems.

NOTES

1. Christopher Manning, "Artificial Intelligence Definitions," Stanford Institute for Human-Centered AI, September 2020, https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf.

2. "Artificial Intelligence—Worldwide," Statista, accessed August 15, 2025, https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide.

3. Nestor Maslej, Loredana Fattorini, Raymond Perrault, et al., *The AI Index 2024 Annual Report*, AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, May 2024, https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf.

4. "Generative AI Could Raise Global GDP by 7%," Goldman Sachs, April 5, 2023, https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html.

5. Maslej, Fattorini, Perrault, et al., *The AI Index 2024 Report*.

6. Jafar Alzubi, Anand Nayyar, and Akshi Kumar, "Machine Learning from Theory to Algorithms: An Overview," *Journal of Physics: Conference Series* 1142, Second National Conference on Computational Intelligence (December 2018), https://doi.org/10.1088/1742-6596/1142/1/012012.

7. "Summary," the Nobel Prize in Physics 2024, the Nobel Prize, October 12, 2024, https://www.nobelprize.org/prizes/physics/2024/summary/.

8. "Summary," the Nobel Prize in Chemistry 2024, the Nobel Prize, October 12, 2024, https://www.nobelprize.org/prizes/chemistry/2024/summary/.

9. Kif Leswing, "Meet the $10,000 Nvidia Chip Powering the Race for A.I.," CNBC, February 23, 2023, https://www.cnbc.com/2023/02/23/nvidias-a100-is-the-10000-chip-powering-the-race-for-ai-.html; Kasper Groes Albin Ludvigsen, "The Carbon Footprint of GPT-4," Medium, July 18, 2023, https://towardsdatascience.com/the-carbon-footprint-of-gpt-4-d6c676eb21ae.

10. Darian Woods and Adrian Ma, hosts, *The Indicator from Planet Money*, podcast, "The Semiconductor Founding Father," NPR, December 21, 2021, 10:14, https://www.npr.org/transcripts/1066548023.

11. Kasper Groes Albin Ludvigsen, "The Carbon Footprint of ChatGPT," *Medium*, December 21, 2022, https://medium.com/data-science/the-carbon-footprint-of-chatgpt-66932314627d.

12. Kasper Groes Albin Ludvigsen, "ChatGPT's Electricity Consumption," Medium, July 12, 2023, https://towardsdatascience.com/chatgpts-electricity-consumption-7873483feac4. Different sources provide somewhat different numbers for the energy cost per query, but they all are in the range of a few watt-hours.

13. "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption," Technology Innovation, EPRI, May 28, 2024, https://www.epri.com/research/products/3002028905.

14. Hope Reese, "A Human-Centered Approach to the AI Revolution," Stanford University Institute for Human-Centered AI, October 17, 2022, https://hai.stanford.edu/news/human-centered-approach-ai-revolution.

15. "Viz.ai Receives New Technology Add-on Payment (NTAP) Renewal for Stroke AI Software from CMS," news release, Viz.ai, August 4, 2021, https://www.viz.ai/news/ntap-renewal-for-stroke-software.

16. Gary Liu, Denise B. Catacutan, Khushi Rathod, et al., "Deep Learning-Guided Discovery of an Antibiotic Targeting *Acinetobacter baumannii*," *Nature Chemical Biology* 19 (2023): 1342–50, https://doi.org/10.1038/s41589-023-01349-8.

17. Albert Haque, Arnold Milstein, and Fei-Fei Li, "Illuminating the Dark Spaces of Healthcare with Ambient Intelligence," *Nature* 585 (2020): 193–202, https://doi.org/10.1038/s41586-020-2669-y.

18. "Innovasea Launches AI-Powered Biomass Camera for Salmon," *The Fish Site*, August 17, 2023, https://thefishsite.com/articles/innovasea-launches-ai-powered-biomass-camera-for-salmon.

19. "Machine Learning in Agriculture: Use Cases and Applications," Itransition.com, February 1, 2023, https://www.itransition.com/machine-learning/agriculture.

20. "J.B. Hunt, Bridgestone and Kodiak Surpass 50,000 Autonomous Long-Haul Trucking Miles in Delivery Collaboration," Kodiak.ai, August 7, 2024, https://kodiak.ai/news/jb-hunt-and-kodiak-collaborate.

21. Steve Lohr, "A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet," *New York Times*, March 19, 2017, https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html.

22. Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, et al., "On the Opportunities and Risks of Foundation Models," preprint, arXiv, Stanford University, July 12, 2022, https://doi.org/10.48550/arXiv.2108.07258.

23. Bommasani, Hudson, Adeli, et al., "On the Opportunities and Risks."

24. Sarah W. Li, Matthew W. Kemp, Susan J. S. Logan, Sebastian E. Illanes, and Mahesh A. Choolani, "ChatGPT Outscored Human Candidates in a Virtual Objective Structured Clinical Examination in Obstetrics and Gynecology," *American Journal of Obstetrics & Gynecology* 229, no. 2 (August 2023): 172.E1–12, https://doi.org/10.1016/j.ajog.2023.04.020.

25. Kent F. Hubert, Kim N. Awa, and Darya L. Zabelina, "The Current State of Artificial Intelligence Generative Language Models Is More Creative Than Humans on Divergent Thinking Tasks," *Scientific Reports* 14, no. 3440 (February 2024), https://doi.org/10.1038/s41598-024-53303-w.

26. Josh Achiam, Steven Adler, Sandhini Agarwal, et al., "GPT-4 Technical Report," preprint, arXiv, March 4, 2024, https://doi.org/10.48550/arXiv.2303.08774.

27. Olawale Salaudeen, Anka Reuel, Ahmed Ahmed, et al., "Measurement to Meaning: A Validity-Centered Framework for AI Evaluation," preprint, arXiv, 2025, https://arxiv.org/pdf/2505.10573.

28. Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of Machine Learning Research* 81, Conference on Fairness, Accountability, and Transparency (February 2018): 1–15, https://www.media.mit.edu/publications/gender-shades-intersectional-accuracy-disparities-in-commercial-gender-classification.

29. Dan Boneh, "Cryptography," chap. 3 in *The Stanford Emerging Technology Review 2025*, Hoover Institution and Stanford School of Engineering (Stanford, CA: Board of Trustees of the Leland Stanford Junior University, 2025), https://doi.org/10.64576/0103.

30. Sayash Kapoor and Arvind Narayanan, "We Looked at 78 Election Deepfakes. Political Misinformation Is Not an AI Problem," Knight First Amendment Institute, December 13, 2024, http://knightcolumbia.org/blog/we-looked-at-78-election-deepfakes-political-misinformation-is-not-an-ai-problem.

31. Neil Perry, Megha Srivastava, Deepak Kumar, and Dan Boneh, "Do Users Write More Insecure Code with AI Assistants?," preprint, arXiv, Cornell University, December 18, 2023, https://doi.org/10.48550/arXiv.2211.03622.

32. Alex Heath, "OpenAI Wants ChatGPT to Be a 'Super Assistant' for Every Part of Your Life," *The Verge*, May 30, 2025, https://www.theverge.com/command-line-newsletter/677705/openai-chatgpt-super-assistant.

33. Zhenghao Yin, Iris Agresti, Giovanni de Felice, et al., "Experimental Quantum-Enhanced Kernel-Based Machine Learning on a Photonic Processor," *Nature Photonics* 19, no. 9 (2025): 1020–27, https://doi.org/10.1038/s41566-025-01682-5.

34. Maja S. Svanberg, Wensu Li, Martin Fleming, Brian C. Goehring, and Neil C. Thompson, "Beyond AI Exposure: Which Tasks Are Cost-Effective to Automate with Computer Vision?," *Future Tech*, Working Paper, January 18, 2024, https://futuretech-site.s3.us-east-2.amazonaws.com/2024-01-18+Beyond_AI_Exposure.pdf.

35. Claire Cain Miller and Courtney Cox, "In Reversal Because of A.I., Office Jobs Are Now More at Risk," *New York Times*, August 24, 2023, https://www.nytimes.com/2023/08/24/upshot/artificial-intelligence-jobs.html.

36. Martin Neil Baily, Erik Brynjolfsson, and Anton Korinek, "Machines of Mind: The Case for an AI-Powered Productivity Boom," Brookings Institution, May 10, 2023, https://www.brookings.edu/articles/machines-of-mind-the-case-for-an-ai-powered-productivity-boom.

37. Pranshu Verma and Gerrit De Vynck, "ChatGPT Took Their Jobs: Now They Walk Dogs and Fix Air Conditioners," *Washington Post*, June 5, 2023, https://www.washingtonpost.com/technology/2023/06/02/ai-taking-jobs/; Challenger, Gray & Christmas, Inc., "Challenger Report," May 2023, https://omscgcinc.wpenginepowered.com/wp-content/uploads/2023/06/The-Challenger-Report-May23.pdf.

38. David Autor, Caroline Chin, Anna M. Salomons, and Bryan Seegmiller, "New Frontiers: The Origins and Content of New Work, 1940–2018," National Bureau of Economic Research, Working Paper 30389, August 2022, https://doi.org/10.3386/w30389.

39. "White House Unveils America's AI Action Plan," The White House, July 23, 2025, https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/.

40. Parliament and Council Regulation 2024/1689 Artificial Intelligence Act, art. 6–7, 2024 O.J. L 2024/1689, https://artificialintelligenceact.eu/section/3-1/.

41. "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023," Department for Science, Innovation, and Technology; Foreign, Commonwealth, and Development Office, Prime Minister's Office, November 1, 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

42. "AI Sovereign Compute Infrastructure Program," Government of Canada, May 5, 2025, https://ised-isde.canada.ca/site/ised/en/ai-sovereign-compute-infrastructure-program; "UK Compute Roadmap," Department for Science, Technology, and Innovation, July 17, 2025, https://www.gov.uk/government/publications/uk-compute-roadmap/uk-compute-roadmap.

43. Eliot Brown and Jin Berber, "SoftBank and OpenAI's $500 Billion AI Project Struggles to Get Off Ground," *Wall Street Journal*, July 21, 2025, https://www.wsj.com/tech/ai/softbank-openai-a3dc57b4.

44. Michael Schuman, "Trump Wants a China Deal That Benefits Him, Not the U.S.," *The Atlantic*, August 14, 2025, https://www.theatlantic.com/ideas/archive/2025/08/china-us-trade-deal-ai-chips/683855/.

45. Final Report, National Security Commission on Artificial Intelligence, March 19, 2021, https://apps.dtic.mil/sti/pdfs/AD1124333.pdf.

46. Charlotte Hill, Charlotte Allen, Tom Perkins, and Harriet Campbell, "Generative AI in the Courts: Getty Images v Stability AI," Penningtons Manches Cooper, February 16, 2024, https://www.penningtonslaw.com/news-publications/latest-news/2024/generative-ai-in-the-courts-getty-images-v-stability-ai.

47. Michael M. Grynbaum and Ryan Mac, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work," *New York Times*, December 27, 2023, https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

48. Blake Brittain, "Music Labels Sue AI Companies Suno, Udio for US Copyright Infringement," Reuters, June 24, 2024, https://www.reuters.com/technology/artificial-intelligence/music-labels-sue-ai-companies-suno-udio-us-copyright-infringement-2024-06-24/.

49. Cade Metz, "Anthropic Agrees to Pay $1.5 Billion to Settle Lawsuit with Book Authors," *New York Times*, September 5, 2025, https://www.nytimes.com/2025/09/05/technology/anthropic-settlement-copyright-ai.html.

50. Blake Brittain, "US Judge Preliminarily Approves $1.5 Billion Anthropic Copyright Settlement," Reuters, September 25, 2025, https://www.reuters.com/sustainability/boards-policy-regulation/us-judge-approves-15-billion-anthropic-copyright-settlement-with-authors-2025-09-25/.

STANFORD EXPERT CONTRIBUTORS

Dr. Fei-Fei Li
SETR Faculty Council, Sequoia Professor in the Computer Science Department, and Professor, by courtesy, of Operations, Information, and Technology at the Graduate School of Business

Dr. Christopher Manning
Thomas M. Siebel Professor of Machine Learning, and Professor of Linguistics and of Computer Science

Anka Reuel
SETR Fellow and PhD Candidate in Computer Science

02

BIOTECHNOLOGY AND SYNTHETIC BIOLOGY

KEY TAKEAWAYS

- Biotechnology is emerging as a general-purpose technology by which anything bioengineers learn to encode in DNA can be grown whenever and wherever needed—essentially enabling the production of a wide range of products through biological processes across multiple sectors.
- The United States is still not executing well on strategies for emerging biotechnology and has relied too heavily on private-sector investment to support foundational work needed to scale and sustain progress.
- Biotechnology is one of the most important areas of technological competition between the United States and China, and China is now leveraging two decades of strategic investment to secure global leadership. Absent swift and ambitious actions, the United States risks biotechnological surprise and a loss of biotechnology sovereignty.

Overview

Biotechnology uses living systems to make products or solve problems. First-generation biotechnology arose over millennia by domesticating and breeding plants and animals[1] for agriculture, food production, companionship, and other purposes.[2] Second-generation biotechnology was launched a half century ago with the invention of recombinant DNA,[3] and it has progressed via techniques including polymerase chain reaction, high-throughput DNA sequencing, and CRISPR gene editing.[4] (DNA is the physical material that encodes biological functions in living systems and is described in more detail later in the chapter.) Both breeding and editing approaches continue to advance, creating ever better tools for sculpting[5] and editing[6] living systems.

Biotechnology products and services are already widely deployed. A 2020 National Academies of Sciences, Engineering, and Medicine report valued the US bioeconomy at around 5 percent of GDP, or more than $950 billion annually.[7] Most applications are in agriculture, medicines, and industrial materials.[8] A 2020 McKinsey & Company report noted that hundreds of biotechnology projects were under development that could add $2 to $4 trillion to the economy.[9] McKinsey's projected doubling of bioeconomic impacts every seven years or so would match biotechnology's economic track record.[10] Its report concluded that, ultimately, biomanufacturing could account for around 60 percent of the global economy's physical inputs.[11] Lowering the cost of biomanufacturing will be essential to realizing such a future.[12]

Synthetic biology continues to emerge as a third wave driving biotechnology, complementing breeding and DNA editing. Synthesis involves putting things together. Synthetic biology explores and adapts concepts from other engineering fields to get better at composing living systems at the molecular, cellular, tissue, and consortia scales. (Consortia refer to biological organization at the level of communities or groups of interacting organisms—typically microbial communities—that function together with division of labor, cooperation, and emergent properties beyond what single cells or tissues can achieve.)

As our ability to compose biology improves, new and more natural modes of biotechnology become possible. For example, leaves on trees do not arrive from factories or central facilities; rather, they grow on trees themselves. Next-generation biotechnology products that operate on a distributed and in situ basis are being explored. For example, a Spanish and British team recently bioengineered plants that can emit light of different colors depending on whether certain viruses are present in the environment around them.[13]

The history of information technology helps in thinking about the emergence of biotechnology. Fifty years ago, computers were mostly industrial, disconnected, and centralized.[14] The emergence of personal computers, packet-switching networks,[15] and programming languages made computing accessible and fun[16] and changed how computer science developed.[17] Biotechnology is poised to experience the same transformation within the next two decades—networked biotechnologies could enable distributed manufacturing resilience, personalized and pervasive biotechnology products, and tools for individual citizens to engage and participate in biotechnology activities.[18]

Key Developments

Analyzing and Understanding Biology with Computing

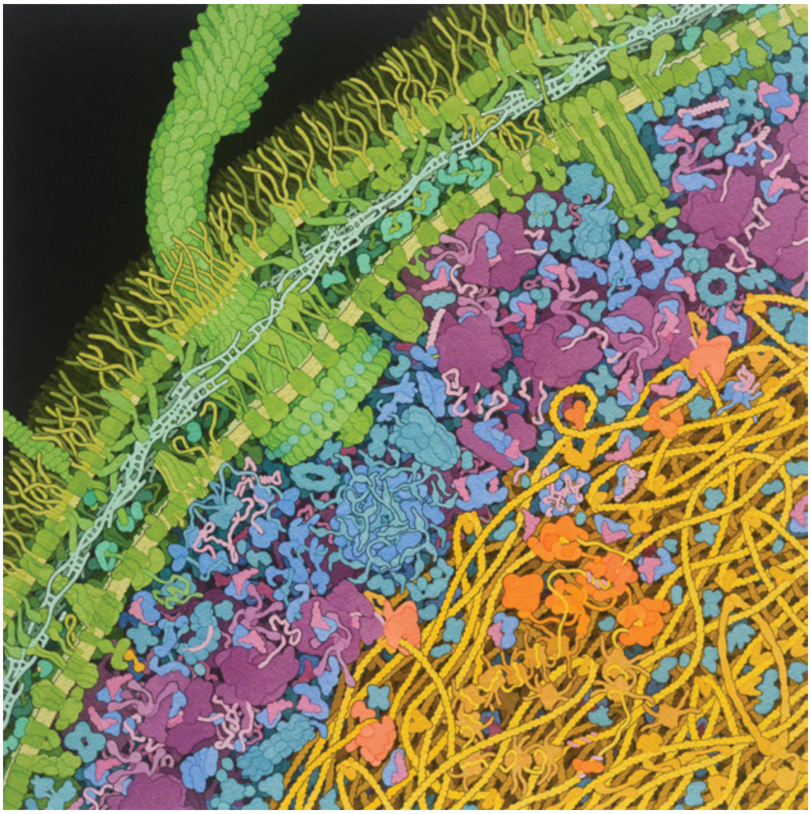
Proteins are molecules that comprise living cells. The shapes of proteins help set their roles and functions. In the 2023 edition of the *Stanford Emerging Technology Review* (*SETR*), we noted how researchers had developed artificial intelligence (AI) methods to estimate the shapes of natural proteins,[19] an accomplishment since recognized with the 2024 Nobel Prize in Chemistry.[20]

Today, AI methods accelerate research by enabling anyone with a computer to estimate the expected shapes of proteins. Researchers can quickly explore what proteins might do without having to run costly real-world experiments. Challenges remain, however. For example, it is still hard to estimate the shapes of proteins that sit within membranes and how certain proteins may change their shape.

Creating AI tools for estimating protein shapes was dependent on decades of laborious experiments by researchers; the actual shapes of enough proteins were needed to train the underlying AI models, and generating these took time. Significant additional data may be needed to develop AI models for understanding protein dynamics, drug-protein interactions, and other life-essential functions.

Cells are the fundamental unit of all living organisms (see figure 2.1). The complexity of cells has long

FIGURE 2.1 Image of a cross section of an *E. coli* cell



Source: David S. Goodsell, RCSB Protein Data Bank, doi: 10.2210/rcsb_pdb/goodsell-gallery-028

motivated researchers to develop computational tools to help make sense of things. There are not yet AI foundation models (described in chapter 1, on artificial intelligence) for representing cells in the same way that there are for representing proteins. Efforts to create "virtual cells" are taking shape, but established measurement methods and computational approaches may be insufficient.[21]

Another emerging approach is to repurpose mathematics used for describing the behavior of materials like toothpaste and ketchup. Such materials, known as "colloidal systems," are mixtures in which particles of one substance are evenly distributed throughout another in a manner similar to how molecules dynamically organize themselves within cells.[22] Recent work that has adapted such mathematical methods to study cells has shown they are capable of representing emergent behaviors that occur at the molecular-to-cellular scale.[23] Combining colloidal models, which model how molecules actually behave within cells, with AI methods is likely the best path toward achieving virtual cells.

Generating and Designing Biology with Computing

Models and software can be further developed to generate novel designs. For example, Chai-2 is an AI model that can design entirely new antibody-based drugs from scratch. (Antibodies are proteins that bind specific target molecules.) In July 2025, Chai-2 achieved a 16 percent success rate in designing antibodies from scratch.[24] That means that 16 percent of the candidates it generates actually result in an effective antibody. Although this percentage may sound low, it compares favorably with traditional experimental lab-based methods that involve screening thousands or millions of candidates with hit rates below 1 percent; it is about one hundred times better than other computational approaches. Chai-2's effectiveness can speed up drug discovery from months or years to weeks. If generalizable, such tools might allow certain medicines to be created faster and more precisely.

Another exciting development is generative biology and genome foundation models, which were introduced in the 2025 edition of *SETR*. Since then, tools like Evo 2 have been released. Evo 2 is described as a "genomic foundation model capable of generalist prediction and design tasks across DNA, RNA, and proteins."[25] A genome foundation model is akin to a large language model trained on natural DNA sequences. When appropriately prompted, these models generate novel DNA sequences. Evo 2 is now capable of generating genomes of viruses, some of which are viable when built and tested in the laboratory.[26]

This capability enables scientists to study viral functions, evolution, and mutations systematically, thereby accelerating discoveries in virology and disease mechanisms. It also facilitates the design of synthetic viruses for beneficial uses like gene therapy, vaccines, and viral delivery systems. Compared to working with natural viruses, it also enables safer, more targeted experimentation.

One bottleneck in generative biology is the limited capacity to interpret or test what a model generates or helps design. No one can yet read entirely novel strings of nucleotides or amino acids and perfectly evaluate the biological function(s) encoded in them. Researchers need faster, better, and larger-scale testing platforms to empirically test how well AI-generated biology actually works.

Distributed Biomanufacturing

The significance of distributed biomanufacturing lies in its flexibility, both in location and timing. For example, because a fermentation process can be established wherever there is access to sugar and electricity, a production site can be set up almost anywhere. The same is true for biomanufacturing processes fed with wood, methane, petroleum, or carbon dioxide (CO_2).

Once established, a biomanufacturing process can respond swiftly to sudden demands, such as those that arise during a pandemic, amid changes in trade policy, or in the case of a war. Such agility can enhance efficiency and revolutionize manufacturing.

One real-world example is the synthetic biology company Antheia, which in early 2024 reported validation of a fermentation-based process for brewing thebaine, a key starting material used in treating opioid overdoses with Narcan.[27] It brewed 116,000-liter batches of bioengineered yeast, with each batch making broth containing a metric ton of thebaine—roughly enough for 100 million Narcan doses.[28] The company's demonstration highlights the potential for on-demand production of critical pharmaceuticals, potentially revolutionizing drug supply chains and improving access to essential medicines.

In 2022, Chinese researchers noted more generally how synthetic biology allows the rewiring of biological systems to support portable, on-site, and on-demand manufacturing of biomolecules.[29] In 2024, Stanford researchers reported on-demand bioproduction of sensors enabling point-of-care health monitoring and detection of environmental hazards aboard the International Space Station.[30] They had already realized many similar demonstrations of distributed biomanufacturing on Earth, ranging from biotechnology education kits to the production of conjugate vaccines.[31]

Such examples demonstrate how biotechnology can be used to make valuable products and services locally. What's happening is a sort of molecular gardening: The energy and material inputs needed to make the biotechnology products are supplied locally, but the process differs from conventional gardening in that bioengineers are programming the genetic instructions for what the biology should make. To fully unlock the power of distributed biomanufacturing, it must also become possible to make the physical DNA encoding genetic programs locally.

Distributed DNA Reading and Writing

DNA is often represented abstractly by its four constituent bases (A, C, T, and G). Unique orderings of these bases encode different biomolecules, which in turn underlie different cellular behaviors and functions.

DNA sequencing (reading of DNA) and synthesis (writing of DNA) are two foundational technologies underlying synthetic biology.[32] Sequencers are machines that determine the precise order of bases in a DNA molecule, effectively converting genetic information from a physical to digital format. Synthesizers generate user-specified digital sequences of A's, C's, T's, and G's, creating physical genetic material from scratch that encodes user-specified sequences, effectively transforming bits into atoms.

If DNA reading and writing tools could themselves be distributed, anyone with an internet connection could upload and download application-specific DNA programs that direct distributed biomanufacturing processes powered by locally available energy and supplied by locally available materials.

In the 1990s, public funding for sequencing the human genome jump-started advances in DNA-sequencing tools by creating significant demand for reading DNA.[33] Private capital and entrepreneurs quickly responded.[34] The Human Genome Project favored development of DNA sequencers that could read billions of bases of DNA as cheaply as possible, resulting in large-format DNA sequencers that were organized in centralized DNA-sequencing factories.[35]

A complementary approach to DNA sequencing has since matured; it allows for individual DNA molecules to be sequenced via tiny pores, or nanopores.[36] UK-based Oxford Nanopore Technologies now offers small DNA sequencers that can be plugged into any computer with a USB port, allowing DNA sequencing to become a distributed technology (figure 2.2).[37]

The market for DNA synthesis has developed slowly over the past forty-five years.[38] Today, most DNA synthesis is carried out via centralized factories.[39] Customers order DNA online and receive materials via express shipping—and it can take days to weeks for the DNA factories to make the DNA molecules. Improvements in commercially available gene-length DNA synthesis services have been modest over the past decade,[40] and in Western countries the services are dependent on private capital.[41]

In June 2025, the United Kingdom's Wellcome Trust launched a Synthetic Human Genome Project via a £10 million seed initiative. The bold goal is to begin to develop tools and infrastructure needed to build synthetic human chromosomes,[42] something that would require significant advances in DNA, gene, and genome-construction technology.

A new generation of companies is also pursuing novel approaches to building DNA—most notably enzymatic DNA synthesis, which uses enzymes and simpler chemical inputs to build DNA.[43] For example, Ansa Biotechnologies is now producing DNA constructs up to 50 kilobases,[44] long enough to construct complete genes, multiple genes, or large genetic elements. This can enable more complex synthetic biology projects, genetic engineering, or even the creation of parts of small viral or bacterial genomes. Enzymatic approaches are compatible with hardware and reagent formats that could potentially enable fast, reliable, and distributed

FIGURE 2.2 Portable DNA sequencers enable biotechnology to become more distributed

Source: Oxford Nanopore Technologies, 2024

DNA synthesis. Significant and sustained investments will be required to make distributed DNA synthesis practical, secure, and accessible.

As another example, researchers in Shanghai have developed a bottom-up DNA synthesis method using DNA origami frameworks—tiny, precisely folded DNA structures—that act like miniature workstations to anchor DNA-writing enzymes at specific positions. Their approach allows ultra-high-density writing of over 500 billion DNA strands per square centimeter, a 10,000-fold improvement over current technologies.[45] If this approach is successfully commercialized, DNA might be written fast enough to support gigabyte-per-second data storage.

Pervasive and Embedded Biotechnologies

The assumption underlying most modern biotechnology products is that they will be carefully contained in steel tanks or used far away from urban populations. However, recent developments in consumer biotechnology products suggest another future. US-based Light Bio, for example, now sells petunia plants bioengineered to emit light (figure 2.3).[46] Light Bio's offering represents an early successful launch of a live consumer biologic, enabling anyone in the United States to source and keep a bioengineered organism for personal use.

In 2024, UK-based Norfolk Plant Sciences first made available to US consumers seeds for its purple tomato, a kind of tomato bioengineered to produce high levels of antioxidants thought to help prevent cancer (see figure 2.4).[47] Stanford faculty bought seeds, and soon bioengineered tomatoes were growing in gardens across campus. Indeed, these tomatoes are available for consumer purchase in a number of grocery stores in the American southeast.[48]

FIGURE 2.3 Light Bio's petunias are bioengineered to emit light

Source: Light Bio Inc.

Another US-based company, ZBiotics, has launched fiber-focused innovations. ZBiotics' Sugar-to-Fiber probiotic drink mix is a genetically engineered microbe that turns dietary sugar into fiber during digestion.[49] This development is emblematic of a next generation of probiotics: They not only supplement the gut microbiome but reshape how the body interacts with food.

An additional category of pervasive and potentially consumer-facing biotechnology involves bioengineering bacteria that live on skin. For example, in 2023, Stanford researchers pioneered the bioengineering of skin microbes to combat skin cancer.[50] Researchers have since expanded such work to enable the eliciting of antigen-specific T cells, which target and eliminate cells infected with viruses and bacteria; these cells also play a role in providing long-term immunological memory.[51] The researchers have even identified specific odorants produced by human-skin microbes whose production could be modulated to reduce mosquito bites.[52]

Meanwhile, CAR-T cell therapy is a new treatment that helps the immune system fight diseases. It was first used for blood cancers but is now being tested for autoimmune diseases and chronic infections.[53] Usually, a patient's own immune cells are taken,

FIGURE 2.4 Norfolk Plant Sciences has bioengineered a more nutritious purple tomato

Source: Norfolk Plant Sciences

changed in a lab to better attack disease, then put back into the body. Researchers are investigating CAR-T cells made from healthy donors, creating off-the-shelf treatments made in advance.[54] These innovations suggest a future where personalized immunotherapies become common, affordable, and widely used medical tools.

Twenty-first-century biotechnologies are increasingly showing up in our homes, on our skin, and in our diets, entering everyday life through familiar channels. These breakthroughs are early indicators of a future in which biology becomes as ubiquitous and integrated in society as electricity or the internet. As noted in a recent hearing of the US-China Economic and Security Review Commission,[55] whichever nation embraces this shift most fully—by "falling in love with biotechnology" not just as science but as part of daily life—will better shape the rules, reap the economic rewards, and lead the next era of bioinnovation.

Over the Horizon

Routinization of Cellular-Scale Engineering

There is no natural cell on Earth that is fully understood. Even well-studied organisms like *E. coli* have genes encoding unknown functions. The simplest and most intensely studied microbes still require more than seventy genes whose functions no researcher understands.[56]

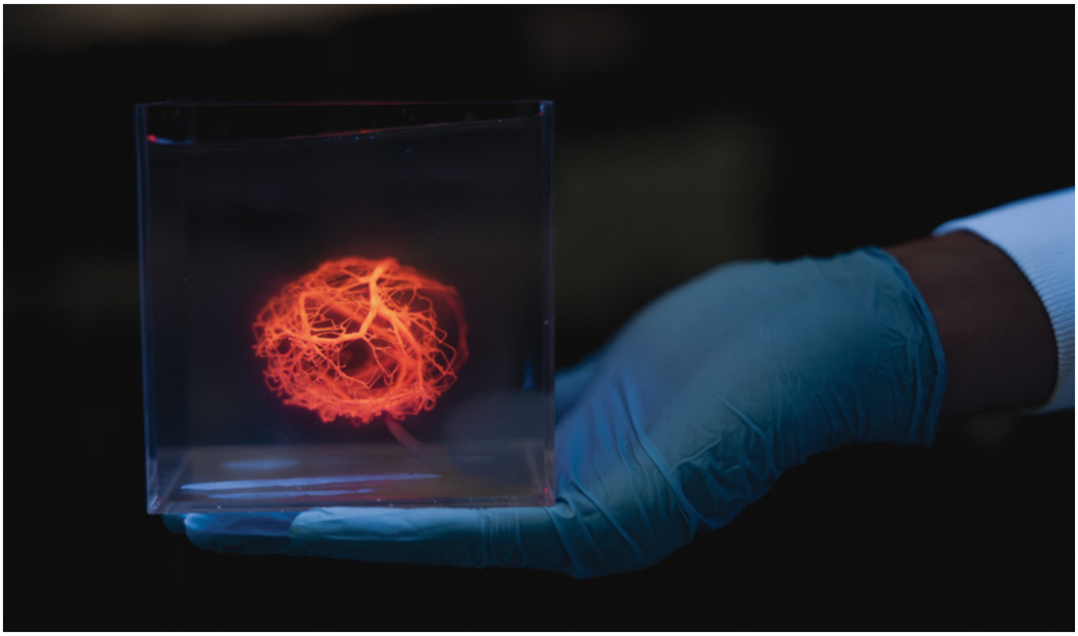
Our ignorance means that biotechnology workflows remain Edisonian at the cellular scale, dependent on tinkering and testing. Bioengineering students are taught "design, build, test, learn"[57] as dogma, with testing requiring many experiments to understand basic phenomenology. Biotechnology projects thus require expert researchers and expensive laboratories while also encountering uncertain budgets and timelines.

Another approach is to build cells starting from only chemically defined molecules. Work based on this approach has rapidly advanced over the past decade, and researchers anticipate soon reporting the first artificial synthetic cells capable of growth, division, and evolution—representing a Sputnik-like milestone for biotechnology. Realizing this potential depends on routinizing bioengineering workflows at the cellular scale to enable "design, build, work" cycles. Such workflows would allow bioengineers to perform relatively minimal empirical validation, focusing primarily on analysis-driven construction of biological artifacts—a hallmark of the engineering rigor found in all modern technologies and essential for future scalable bioengineering.[58]

Constructing simple life from scratch will enable the transcendence of constraints on Earth's lifeforms[59]—organisms limited by lineage and requirements of reproduction and evolvability. A next level of biotechnologies will be unlocked, providing a perch from which to access everything biology can become. More practically, gaining the capacity to make the engineering of cellular-scale systems

Researchers anticipate soon reporting the first artificial synthetic cells capable of growth, division, and evolution—representing a Sputnik-like milestone for biotechnology.

FIGURE 2.5 A model of a vascular tree printed using a 3-D bioprinter



Source: Andrew Brodhead, *Stanford Report*

routine will enable the development and deployment of biotechnologies on a much more reliable basis.

Printing Tissues on Demand

Tissue printing uses living cells as inks to construct tissue-like structures. Early methods relied on sparse inks around one hundred times less dense than natural tissue. Recent breakthroughs are enabling printing with about 200 million cells per milliliter,[60] a density approaching that found in live organs.

At Stanford, Mark A. Skylar-Scott and his team have developed a method called SWIFT (sacrificial writing into functional tissue) that helps create living heart tissue in the lab. Starting with hundreds of thousands of tiny clusters of special stem cells and mixing them into a soft paste, they "print" tiny channels inside this mixture, similar to making blood vessels. This allows oxygen and nutrients to flow through the tissue and helps the cells survive, join together, and even start beating like a real heart (figure 2.5).

Making enough cells remains a key challenge. Skylar-Scott's team can generate billions of heart-specific cells every two weeks via automated bioreactors.[61] Further increases in cell production combined with organ-scale design will be needed to print implant-ready tissues.[62]

Such developments reflect a broader shift in the field of tissue engineering and printing—from simple cell sheets to dense, vascularized, and physiologically active tissue systems. Once confined to small-scale tissue patches or skin grafts, the field is now generating, among other things, tissue that supports perfusion, which covers flowing blood or fluid through blood vessels to oxygenate and feed tissues and organs; functional integration (i.e., an integration of multiple components in a tissue or other biological construct that performs desired biological functions in a reliable and functional manner); and long-term viability of engineered tissues. The dreams of building whole-tissue or full-organ fabrication are shifting from speculation to matters of engineering rigor, scale, investment, and translation.

Electrobiosynthesis

Carbon is central to life. Currently, photosynthesis captures CO_2 from the atmosphere to produce organic carbon molecules. Recent thinking, however, suggests that electricity could be used to fix carbon directly from the air to create organic molecules that could be fed to microbes—a process that may become known as electrobiosynthesis, or, more simply, as "eBio." Capturing carbon in this way could be an order of magnitude more efficient from a land-use perspective than traditional agriculture.[63]

The idea is to engineer a parallel carbon cycle that starts with air and electricity, perhaps generated via solar panels, to create organic molecules that power bioproduction processes. For example, in 2024, Stanford researchers reported the creation of a system that combines electrochemistry with biological processes to transform simple carbon compounds into a key organic molecule called acetyl-CoA (acetyl coenzyme A), which is present in all living things and acts as a building block for other molecules within cells.[64]

Although eBio is a very immature technology, its potential significance and impacts are hard to overstate. For example, surplus power from large-scale renewable energy generation might directly produce biomolecules such as proteins and cellulose without requiring massive conventional battery banks to store energy that cannot be used immediately. The development of eBio could also enable bioproduction in places where soils are poor, water is scarce, or climate and weather are too uncertain.

Ultimately, eBio could increase how much humanity can make in partnership with biology. We would be constrained only by how much energy we can generate for such purposes. This approach could significantly reduce the land and water requirements for biomass production, potentially alleviating pressure on agricultural resources and offering a more sustainable path for biomanufacturing.

Biology as a General-Purpose Technology

Biotechnology is now used to make medicines, foods, and a relatively narrow range of materials (e.g., sustainable carpet fibers). However, anything whose biosynthesis engineers can learn to encode in DNA could be grown using biology. Examples from nature highlight the potential here: Some bacteria naturally grow arrays of tiny magnets,[65] while select sea sponges grow glass filaments like fiber-optic cables.[66] These bio-made magnets and filaments form under ambient conditions through naturally sustainable processes and can be more robust than conventional alternatives. Such examples fuel calls for biology to be recognized as a general-purpose technology that, with appropriate vision and leadership, could become the foundation of a much more resilient manufacturing base.[67]

As one example, in 2018 the Semiconductor Research Corporation (SRC) offered an ambitious twenty-year synthetic biology road map toward growing computers.[68] SRC's first proposed step was to develop DNA for archival data storage.[69] In 2024 the Hoover Library & Archives partnered with Twist Bioscience to encode a digital copy of the telegram from President Hoover founding his namesake institution within synthetic DNA contained in a tiny ampule (see figure 2.6). Made in this way, the DNA serves as a data storage medium whose digital contents must be recovered via DNA sequencing. In April 2025, the Library of Congress requested proposals to store 1 terabyte of data in synthetic DNA. The intention is to provide "both a functional and artistic display" of some of the nation's digital treasures in celebration "of the 250th anniversary of the signing of the Declaration of Independence."[70]

Many other things must be made real to ever grow computers. Scattered progress is happening: In 2024 researchers in California reported using a synthetic

FIGURE 2.6 DNA is used as a storage medium for a digital copy of Herbert Hoover's telegram founding his namesake institution

monolayer of DNA origami to assemble and study solid-state spin qubits for quantum sensing applications. (Qubits and quantum sensing are described further in chapter 7, on quantum technologies.)[71]

Another example is the 2011 US Navy program Application of Synthetic Biological Techniques for Energetic Materials.[72] This program explored the ability to brew propellants and explosives—an ability that could enable any nation to create more resilient supply chains for key military materials. For example, a distributed and resilient biomanufacturing network could help NATO members meet their Article 3 obligations related to supply chain resilience.[73]

Unlocking biology as a general-purpose technology by learning to grow computers, energetics, and many other things might require $100 billion in well-managed foundational research over a twenty-year period. As yet, no such coordinated effort is underway.

Policy Issues

Getting Private and Public Investments Right

Many first-generation synthetic biology companies continue to struggle.[74] Billions of dollars of private capital have been lost in biotechnology investments made in the United States alone over the past two decades. One perspective is that these early big bets were simply too early.[75] The hope is that small and scrappy efforts will eventually find their way to success. However, an immediate issue is that many sources of private capital that could support next-generation biotechnology companies are now shut off because those prior bets were unsuccessful; this adds sector-specific headwinds to the general challenges that face young, innovative businesses.

Another perspective is that America has relied too heavily on the private sector to invent, advance, and deploy emerging biotechnologies.[76] Because many private investors expect foundational advances and platforms to quickly generate and sustain revenue growth in order to justify increased valuations and further funding, young biotechnology companies go to market too early and experience a pattern of repeated failures. A June 2025 congressional hearing explored breaking this cycle via smart and sustained public investments in foundational bioengineering research, including tools for measuring, modeling, and making biology and public-benefit research platforms.[77]

Safety and Security Concerns

New organisms can raise concerns about how they might interact with natural or human environments. Bioengineered organisms might disrupt local ecosystems. Malicious actors could create organisms harmful to people or environments.[78]

A specific recent concern is the potential construction of mirror life and mirror microorganisms. Mirror life—entirely hypothetical today—is made of biological molecules that are mirror images of those found in natural life on Earth. For example, all known DNA is left-handed, referring to the spiral shape of DNA. Mirror DNA, which would be found in mirror cells, would be right-handed. All biological molecules found on Earth—DNA, sugars, fats, and proteins—have handedness, and mirror versions of these molecules would have the opposite handedness.

Mirror life is the natural end point of research on mirror-image biology, which has potential value in a variety of applications. For example, mirror-image molecules might help therapeutic drugs resist undesirable natural enzymatic degradation in the human body. Mirror cells could be the most economical way to produce such molecules.

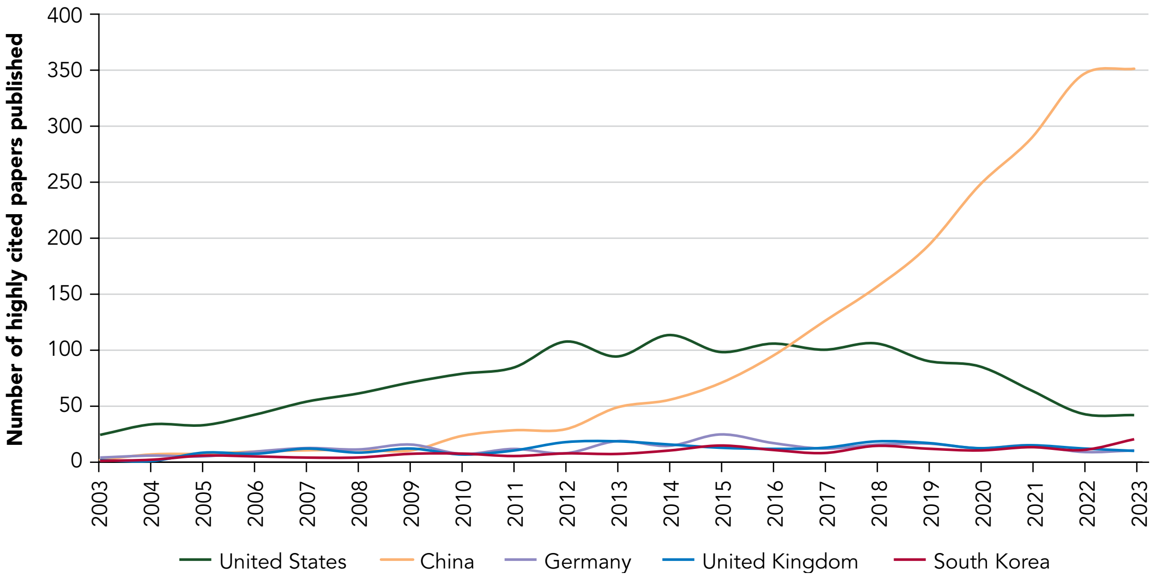
However, if they escaped into the environment, such organisms might not be readily recognized by the immune systems of plants and animals, including humans. The increasing risk of someone making mirror microbes has resulted in calls for outright bans and prohibitions on such work.[79]

Reflecting such types of concern, experts in the life sciences convened in February 2025 for the Spirit of Asilomar and the Future of Biotechnology summit. It took place on the fiftieth anniversary of the original Asilomar conference, which had convened to address the potential biohazards of recombinant DNA. The 1975 conference ultimately resulted in the establishment of voluntary guidelines and safety principles for the responsible conduct of genetic engineering research and set an important precedent for self-regulation in biotechnology.

The 2025 Spirit of Asilomar was organized around discussions of several themes,[80] including pathogens research and biological weapons, AI and biotechnology, synthetic cells, biotechnologies beyond conventional containment, and the framing of biotechnology's futures. The conference resulted in the publication of twenty-seven entreaties—public recommendations and calls for dialogue spanning these themes.[81] The 2025 conference organizers hope that these entreaties will constitute an important point of departure for proactive, forward-looking governance that addresses both known risks and emerging challenges at the intersection of synthetic biology, AI, and global biosecurity.

Additionally, concerns about biosafety and biosecurity have stimulated interest in a variety of control measures to ensure appropriate use of biotechnology and strengthen the governance of pathogen-related research.[82] For example, a rapid increase in the deployment of BSL-3 and BSL-4 laboratories (biological laboratories with the highest biosafety levels and thus the most stringent safety and security measures) reflects heightened attention to biosafety and biosecurity needs as more researchers work with higher-risk pathogens and synthetic biology tools.

FIGURE 2.7 China is outpacing the United States in publishing highly cited research papers on synthetic biology



Source: Adapted from Australian Strategic Policy Institute, Critical Technology Tracker, based on "Appendix 2: Detailed Methodology," in Jennifer Wong-Leung, Stephan Robin, and Danielle Cave, ASPI's Two-Decade Critical Technology Tracker, August 2024

Ethical Considerations

Different religious traditions may have different stances toward life and whether the engineering of new life-forms violates any of their basic precepts. In the words of a report published by the Woodrow Wilson International Center for Scholars, such concerns involve "the possibility of harm to deeply held (if sometimes hard to articulate) views about what is right or good, including . . . the appropriate relationship of humans to themselves and the natural world."[83] Just because something might or can be done does not mean that it should be done.

As Drew Endy (the SETR faculty member for biotechnology) and Laurie Zoloth note, "The narrative of creation of the human is the central narrative for many religious communities. To create a human genome from scratch would be an enormous moral gesture whose consequences should not be framed initially on the advice of lawyers and regulators alone."[84]

Global Competition

The United States and other nations continue to develop, advance, and refine strategies for biotechnology, biomanufacturing innovation, biosecurity, and the overall bioeconomy. For example, the United States' Congressional National Security Commission on Emerging Biotechnology has now published its final report,[85] calling for bold investments in emerging biotechnology domestically.

From a competition perspective, many have been sounding the alarm that China risks outpacing the United States and Europe in established and emerging biotechnology, from research to innovation to commercialization and full-scale manufacturing.[86] For example, in 2023 researchers in China published nearly 350 papers that ranked among the top 10 percent most-cited papers on synthetic biology. This is compared to 41 such papers in the United States (see figure 2.7). As another example, licensing of novel drugs has seen a dramatic shift, with China-based firms reporting an almost twentyfold increase in licensing deals over the past decade.[87]

If the United States and Europe fail to match China's all-of-nation support for biotechnology, then neither will maintain biotechnology leadership or sovereignty going forward. Absent dramatic action, such failure seems a most likely outcome; plaintive warnings or calls for action from researchers in the West simply help spur greater investments from Beijing.

NOTES

1. Jared Diamond, "Evolution, Consequences, and Future of Plant and Animal Domestication," *Nature* 418 (August 2020): 700–707, https://doi.org/10.1038/nature01019.

2. Freeman Dyson, "Our Biotech Future," *New York Review of Books* 54, no. 12 (July 2007): 4–7, https://www.nybooks.com/articles/2007/07/19/our-biotech-future/.

3. Tim Beardsley, "Biotechnology: Cohen-Boyer Patent Finally Confirmed," *Nature* 311, no. 5981 (September 1984): 3, https://doi.org/10.1038/311003a0.

4. Irina Gostimskaya, "CRISPR-Cas9: A History of Its Discovery and Ethical Considerations of Its Use in Genome Editing," *Biochemistry* (Moscow) 87, no. 8 (August 2022): 777–78, https://doi.org/10.1134/S0006297922080090. CRISPR is an acronym for clustered regularly interspaced short palindromic repeats.

5. Frances H. Arnold, "Innovation by Evolution: Bringing New Chemistry to Life," December 8, 2018, Stockholm University, Sweden, PDF, https://www.nobelprize.org/prizes/chemistry/2018/arnold/lecture/.

6. "Genetic Scissors: A Tool for Rewriting the Code of Life," news release, the Nobel Prize, October 7, 2020, https://www.nobelprize.org/prizes/chemistry/2020/press-release/.

7. National Academies of Sciences, Engineering, and Medicine, "Safeguarding the Bioeconomy" (National Academies Press, 2020), 73, https://doi.org/10.17226/25525.

8. Planetary Technologies, "Bioeconomy Dashboard," last modified 2023, https://www.planetarytech.earth/bioeconomy-dashboard-1.

9. Michael Chui, Matthias Evers, James Manyika, Alice Zheng, and Travers Nisbet, "The Bio Revolution: Innovations Transforming Economies, Societies, and Our Lives," McKinsey & Company, May 13, 2020, https://www.mckinsey.com/industries/life-sciences/our-insights/the-bio-revolution-innovations-transforming-economies-societies-and-our-lives.

10. Robert Carlson, "Estimating the Biotech Sector's Contribution to the US Economy," *Nature Biotechnology* 34, no. 3 (March 2016): 247–55, https://doi.org/10.1038/nbt.3491.

11. Chui, Evers, Manyika, Zheng, and Nisbet, "The Bio Revolution."

12. Jennifer Hennigan, Phillip Wagner, Chris Burk, et al., "A Technoeconomic Evaluation of the Potential of Industrial Biotechnology for the Competitive Production of Commodity and Bulk Chemicals," preprint, ChemRxiv, November 16, 2020, https://doi.org/10.26434/chemrxiv.13238996.v1.

13. Camilo Calvache, Marta Rodriguez-Rodriguez, Victor Vazquez-Vilriales, et al., "Bioluminescent Sentinel Plants Enable Autonomous Diagnostics of Viral Infections," preprint, bioRxiv, January 1, 2025, 2025.08.05.668616, https://doi.org/10.1101/2025.08.05.668616.

14. "IBM Mainframe," Wikipedia, last modified July 20, 2024, https://en.wikipedia.org/wiki/IBM_mainframe.

15. Vinton G. Cerf and Robert E. Kahn, "A Protocol for Packet Network Intercommunication," *IEEE Transactions on Communications* 22, no. 5 (May 1974): 637–48, https://doi.org/10.1109/TCOM.1974.1092259.

16. *Whole Earth Epilog* (Penguin Books, 1974), https://archive.org/details/wholeearthepilog00unse.

17. "History of Personal Computers," Wikipedia, last modified August 19, 2024, https://en.wikipedia.org/wiki/History_of_personal_computers.

18. Callie R. Chappell, Ana Paulina Quiroz, David Sun Kong, and Drew Endy, "Creating a Popular Foundation for the Bio-Age," *Issues in Science and Technology* 41, no. 4 (Summer 2025): 60–63, https://doi.org/10.58875/LCHC2652.

19. Ewen Callaway, "'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein," *Nature*, July 29, 2023, https://www.nature.com/articles/d41586-022-02083-2.

20. "Summary," the Nobel Prize in Chemistry 2024, the Nobel Prize, October 12, 2024, https://www.nobelprize.org/prizes/chemistry/2024/summary/.

21. Elliot Hershberg, "What Are Virtual Cells?," The Century of Biology, June 29, 2025, https://centuryofbio.com/p/virtual-cell.

22. Akshay J. Maheshwari, Alp M. Sunol, Emma Gonzalez, Drew Endy, and Roseanna N. Zia, "Colloidal Hydrodynamics of Biological Cells: A Frontier Spanning Two Fields," *Physical Review Fluids* 4, no. 11 (2019): 110506, https://doi.org/10.1103/PhysRevFluids.4.110506.

23. Akshay J. Maheshwari, Jonathan Calles, Sean K. Waterton, and Drew Endy, "Engineering tRNA Abundances for Synthetic Cellular Systems," *Nature Communications* 14, no. 1 (2023): 4594, https://doi.org/10.1038/s41467-023-40199-9; Akshay J. Maheshwari, Alp M. Sunol, Emma Gonzalez, Drew Endy, and Roseanna N. Zia, "Colloidal Physics Modeling Reveals How Per-Ribosome Productivity Increases with Growth Rate in Escherichia Coli," *mBio* 14, no. 1 (2022): e02865–22, https://doi.org/10.1128/mbio.02865-22.

24. Chai Discovery Team et al., "Zero-Shot Antibody Design in a 24-Well Plate," preprint, bioRxiv, July 6, 2025, https://doi.org/10.1101/2025.07.05.663018.

25. "Evo 2: DNA Foundation Model," Arc Institute, 2024, https://arcinstitute.org/tools/evo; Eric Nguyen, Michael Poli, Matthew G. Durrant, et al., "Sequence Modeling and Design from Molecular to Genome Scale with Evo," preprint, bioRxiv, March 6, 2024, https://doi.org/10.1101/2024.02.27.582234.

26. Samuel H. King, Claudia L. Driscoll, David B. Li, et al., "Generative Design of Novel Bacteriophages with Genome Language Models," preprint, bioRxiv, January 1, 2025, 2025.09.12.675911, https://doi.org/10.1101/2025.09.12.675911.

27. "Antheia Completes Successful Product Validation," Antheia, January 8, 2024, https://antheia.bio/antheia-completes-successful-product-validation/.

28. "U.S. Secretary of State Antony Blinken Visits Antheia to Discuss Biotechnology Innovation," Antheia, May 30, 2024,

https://antheia.bio/u-s-secretary-of-state-antony-j-blinken-visits-antheia-to-discuss-biotechnology-innovation/.

29. Chenwang Tang, Lin Wang, Lei Zang, Qing Wang, Dianpeng Qi, and Zhuojun Dai, "On-Demand Biomanufacturing Through Synthetic Biology Approach," *Materials Today Bio* 18, no. 100518 (February 2023), https://doi.org/10.1016/j.mtbio.2022.100518.

30. Selin Kocalar, Bess M. Miller, Ally Huang, et al., "Validation of Cell-Free Protein Synthesis Aboard the International Space Station," *ACS Synthetic Biology* 13, no. 3 (March 2024): 942–50, https://doi.org/10.1021/acssynbio.3c00733.

31. Jessica C. Stark, Thapakorn Jaroentomeechai, Tyler D. Moeller, et al., "On-Demand Biomanufacturing of Protective Conjugate Vaccines," *Science Advances* 7, no. 6 (February 2021): eabe9444, https://doi.org/10.1126/sciadv.abe9444; Jessica C. Stark, Ally Huang, Peter Q. Nguyen, et al., "BioBits™ Bright: A Fluorescent Synthetic Biology Education Kit," *Science Advances* 4, no. 8 (August 2018): eaat5107, https://doi.org/10.1126/sciadv.aat5107.

32. National Security Commission on Emerging Biotechnology, "DNA: Reading, Writing, and Editing," February 2024, https://www.biotech.senate.gov/press-releases/dna-reading-writing-and-editing/.

33. "The Human Genome Project," National Human Genome Research Institute, National Institutes of Health, last modified May 14, 2024, https://www.genome.gov/human-genome-project.

34. Kristen Philipkoski, "Celera Wins Genome Race," *Wired*, April 6, 2000, https://www.wired.com/2000/04/celera-wins-genome-race/.

35. James M. Heather and Benjamin Chain, "The Sequence of Sequencers: The History of Sequencing DNA," *Genomics* 107, no. 1 (January 2016): 1–8, https://doi.org/10.1016/j.ygeno.2015.11.003.

36. Davied Deamer, Mark Akeson, and Daniel Branton, "Three Decades of Nanopore Sequencing," *Nature Biotechnology* 34 (May 2016): 518–24, https://doi.org/10.1038/nbt.3423.

37. "Company History," Oxford Nanopore Technologies, accessed September 3, 2024, https://nanoporetech.com/about/history.

38. Marvin H. Caruthers, "The Chemical Synthesis of DNA/RNA: Our Gift to Science," *Journal of Biological Chemistry* 288, no. 2 (December 2012): 1420–24, https://doi.org/10.1074/jbc.X112.442855.

39. "Integrated DNA Technologies Invests in New U.S. Synthetic Biology Manufacturing Facility," Business Wire, May 28, 2024, https://www.businesswire.com/news/home/20240528727951/en/Integrated-DNA-Technologies-Invests-in-New-U.S.-Synthetic-Biology-Manufacturing-Facility.

40. Planetary Technologies, "Bioeconomy Dashboard."

41. "Twist Bioscience Expands Gene Offering with Long Gene Fragments up to 5.0kb," Business Wire, August 8, 2024, https://www.businesswire.com/news/home/20240808863612/en/Twist-Bioscience-Expands-Gene-Offering-With-Long-Gene-Fragments-up-to-5.0kb.

42. "New Project to Pioneer the Principles of Human Genome Synthesis," Wellcome Trust, June 26, 2025, https://wellcome.org/news/new-project-pioneer-principles-human-genome-synthesis.

43. MaryAnn Labant, "Enzymatic DNA Synthesis: Shorter Waits, Longer Strands," *Genetic Engineering & Biotechnology News*, July 1, 2024, https://www.genengnews.com/topics/omics/enzymatic-dna-synthesis-shorter-waits-longer-strands.

44. "Ansa Biotechnologies Announces Successful de novo Synthesis of World's Longest Oligonucleotide at 1005 Bases," CRISPR Medicine News, March 9, 2023, https://crisprmedicinenews.com/press-release-service/card/ansa-biotechnologies-announces-successful-de-novo-synthesis-of-worlds-longest-oligonucleotide-at-10-1/.

45. Chunhong Li, Yishakejiang Saimaiti, Min Li, et al., "DNA Framework Array Enables Ultra-High Throughput DNA Synthesis," preprint, bioRxiv, May 31, 2025, https://doi.org/10.1101/2025.05.30.657018.

46. "Light Bio," Light Bio, accessed September 3, 2024, https://www.light.bio.

47. Sasa Woodruff, "Gardeners Can Now Grow a Genetically Modified Purple Tomato Made with Snapdragon DNA," NPR, February 6, 2024, https://www.npr.org/sections/health-shots/2024/02/06/1228868005/purple-tomato-gmo-gardeners.

48. "Empress Limited Edition Tomato," Norfolk Healthy Produce, accessed September 3, 2024, https://www.norfolkhealthyproduce.com.

49. "ZBiotics Sugar-to-Fiber Probiotic Drink Mix," Prepared Foods, October 12, 2024, https://www.preparedfoods.com/articles/129812-zbiotics-sugar-to-fiber-probiotic-drink-mix.

50. Hadley Leggett, "Researchers Use Skin-Colonizing Bacteria to Create a Topical Cancer Therapy in Mice," Stanford Medicine, April 12, 2023, https://med.stanford.edu/news/all-news/2023/04/cancer-bacteria.html.

51. Michael A. Fischbach, Kazuki Nagashima, Yiyin E. Chen, and Djenet Bousbaine, "Bacteria-Engineered to Elicit Antigen-Specific T Cells," US Patent 2024/0024380 A1, filed December 22, 2021, and issued January 25, 2024.

52. Iliano V. Coutinho-Abreu, Omid Jamshidi, Robyn Raban, Katayoon Atabakhsh, Joseph A. Merriman, and Omar S. Akbari, "Identification of Human Skin Microbiome Odorants that Manipulate Mosquito Landing Behavior," *Scientific Reports* 14, no. 1631 (January 2024), https://doi.org/10.1038/s41598-023-50182-5.

53. Yaojie Kong, Jingyao Li, Xueyao Zhao, et al., "CAR-T Cell Therapy: Developments, Challenges and Expanded Applications from Cancer to Autoimmunity," *Frontiers in Immunology* 15 (2024), https://doi.org/10.3389/fimmu.2024.1519671.

54. "CAR T Cells: Engineering Patients' Immune Cells to Treat Their Cancers," National Cancer Institute, February 26, 2025, https://www.cancer.gov/about-cancer/treatment/research/car-t-cells.

55. Drew Endy, "STRANGE COMPETITION. A Statement of Evidence Written in 2025," testimony presented before the US-China Economic and Security Review Commission, hearing on "Made in China 2025—Who Is Winning?," February 6, 2025.

56. John I. Glass, Chuck Merryman, Kim S. Wise, Clyde A. Hutchison III, and Hamilton O. Smith, "Minimal Cells—Real and Imagined," *Cold Spring Harbor Perspectives in Biology* 9, no. 12 (December 2017): a023861, https://doi.org/10.1101/cshperspect.a023861.

57. Shohei Kitano, Ciai Lin, Jee Loon Foo, and Matthew Wook Chang, "Synthetic Biology: Learning the Way Toward High-Precision Biological Design," *PLOS Biology* 21, no. 4 (April 2023): e3002116, https://doi.org/10.1371/journal.pbio.3002116.

58. A useful description of the contrast between "design-build-test-learn" and "design-build-work" can be found at https://centuryofbio.com/p/design-build-work.

59. Drew Endy, "Upwelling," *Original Syn*, blog, October 28, 2020, https://blog.originalsyn.bio/2020/10/upwelling.html.

60. Mark A. Skylar-Scott, Sebastien G. M. Uzel, Lucy L. Nam, et al., "Biomanufacturing of Organ-Specific Tissues with High Cellular

Density and Embedded Vascular Channels," *Science Advances* 5, no. 9 (September 6, 2019): eaaw2459, https://doi.org/10.1126/sciadv.aaw2459.

61. "Moonshot Effort Aims to Bioprint a Human Heart and Implant It in Pig," news release, Stanford News, September 23, 2023, https://news-archive.stanford.edu/press-releases/2023/09/28/moonshot-effort-eart-implant-pig.

62. Zachary A. Sexton, Dominic Rütsche, Jessica E. Herrmann, et al., "Rapid Model-Guided Design of Organ-Scale Synthetic Vasculature for Biomanufacturing," *Science* 388, no. 6752 (June 12, 2025): 1198–1204, https://doi.org/10.1126/science.adj6152.

63. Dorian Leger, "Photovoltaic-Driven Microbial Protein Production Can Use Land and Sunlight More Efficiently than Conventional Crops," *Proceedings of the National Academy of Sciences USA* 118, no. 26 (June 2021): e2015025118, https://doi.org/10.1073/pnas.2015025118; Emiliano Bellini, "Solar-Powered Large Scale Microbial Food Production," *pv Magazine*, August 3, 2021, https://www.pv-magazine.com/2021/08/03/solar-powered-large-scale-microbial-food-production.

64. Grant M. Landwehr, Bastian Vogeli, Cong Tian, et al., "A Synthetic Cell-Free Pathway for Biocatalytic Upgrading of One-Carbon Substrates," preprint, bioRxiv, August 8, 2024, https://doi.org/10.1101/2024.08.08.607227.

65. Pranami Goswami, Kuang He, Jinhua Li, Yongxin Pan, Andrew P. Roberts, and Wei Lin, "Magnetotactic Bacteria and Magnetofossils: Ecology, Evolution, and Environmental Implications," *NPJ Biofilms and Microbiomes* 8, no. 43 (June 2022), https://doi.org/10.1038/s41522-022-00304-0.

66. Sarah Graham, "Sea Sponge Inspires Better Fiber-Optic Cables," *Scientific American*, August 21, 2003, https://www.scientificamerican.com/article/sea-sponge-inspires-bette/.

67. Abigail Kukura, PJ Maykish, David Lin, et al., "National Action Plan for U.S. Leadership in Biotechnology," Special Competitive Studies Project, April 12, 2023, 1, https://www.scsp.ai/wp-content/uploads/2023/04/National-Action-Plan-for-U.S.-Leadership-in-Biotechnology.pdf.

68. "SemiSynBio Consortium and Roadmap Development," Semiconductor Research Corporation, accessed September 3, 2024, https://www.src.org/program/grc/semisynbio/semisynbio-consortium-roadmap.

69. "DNA Data Storage Alliance," DNA Data Storage Alliance, accessed September 3, 2024, https://dnastoragealliance.org.

70. "Synthetic DNA Data Storage—Library of Congress Request for Information," Sam.gov, 2025, https://sam.gov/opp/adc4a2ac1f2445de83eacfc62a8e75dd/view.

71. Zhiran Zhang, "Probing and Engineering the Environment of Near-Surface Nitrogen-Vacancy Centers in Diamond for Quantum Sensing and Simulation" (PhD diss., University of California–Santa Barbara, 2024).

72. "Synthetic Biological Techniques for Energetic Materials," Strategic Environmental Research and Development Program, Environmental Security Technology Certification Program, accessed September 3, 2024, https://serdp-estcp.mil/newsitems/details/ac878993-2005-4182-948a-f63c95668499/synthetic-biological-techniques-for-energetic-materials.

73. "Resilience in NATO," NATO Allied Command Transformation, December 15, 2023, https://www.act.nato.int/article/resilience-in-nato/.

74 ."Biotech Firm Amyris Files for Bankruptcy in US," Reuters, August 10, 2023, https://www.reuters.com/business/biotech-firm-amyris-files-bankruptcy-us-2023-08-10; Amy Feldman and Angel Au-Yeung, "The Inside Story of How SoftBank-Backed Zymergen Imploded Four Months After Its $3 Billion IPO," *Forbes*, October 13, 2021, https://www.forbes.com/sites/amyfeldman/2021/10/13/the-inside-story-of-how-softbank-backed-zymergen-imploded-four-months-after-its-3-billion-ipo.

75. Robert F. Service, "Synthetic Biology, Once Hailed as a Moneymaker, Meets Tough Times," *Science*, August 22, 2024, https://www.science.org/content/article/synthetic-biology-once-hailed-moneymaker-meets-tough-times.

76. Drew Endy, "Funding Biotechnology's Foundations to Fuel Private Sector Innovations," testimony, *Pursuing the Golden Age of Innovation: Strategic Priorities in Biotechnology*, a joint meeting of the Research and Technology and Energy Subcommittees of the Committee on Science, Space, and Technology of the US House of Representatives, June 5, 2025, https://www.hoover.org/research/joint-research-technology-and-energy-subcommittee-hearing-pursuing-golden-age-innovation-0.

77. *Pursuing the Golden Age of Innovation: Strategic Priorities in Biotechnology*, a joint meeting of the Research and Technology and Energy Subcommittees of the Committee on Science, Space, and Technology of the US House of Representatives, June 5, 2025, https://science.house.gov/2025/6/pursuing-the-golden-age-of-innovation-strategic-priorities-in-biotechnology.

78. For example, polio, horsepox, SARS-CoV-2, and the Spanish flu virus have been synthesized from scratch in laboratories. See, respectively, Jeronimo Cello, Aniko V. Paul, and Eckard Wimmer, "Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template," *Science* 297, no. 5583 (2002): 1016–18, https://www.science.org/doi/10.1126/science.1072266; Ryan S. Noyce, Seth Lederman, David H. Evans, "Construction of an Infectious Horsepox Virus Vaccine from Chemically Synthesized DNA Fragments," *PLOS ONE* 13, no. 1 (2018): e0188453, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0188453; Tran Thi Nhu Thao, Fabien Labroussaa, Nadine Ebert, et al., "Rapid Reconstruction of SARS-CoV-2 Using a Synthetic Genomics Platform," *Nature* 582 (May 2020): 561–65, https://www.nature.com/articles/s41586-020-2294-9; Terrence M. Tumpey, Christopher F. Basler, Patricia V. Aguilar, et al., "Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus," *Science* 310, no. 5745 (2005): 77–80, https://www.science.org/doi/10.1126/science.1119392.

79. Katarzyna P. Adamala, Deepa Agashe, Yasmine Belkaid, et al., "Confronting Risks of Mirror Life," *Science* 386, no. 6728 (2024): 1351–53, https://doi.org/10.1126/science.ads9158.

80. "Summit Themes," Spirit of Asilomar, 2025, https://www.spiritofasilomar.org/program/summit-themes.

81. "The Spirit of Asilomar and the Future of Biotechnology," Rice Research Repository, 2025, https://repository.rice.edu/communities/4825def5-159e-4969-ae67-b2faf0c3b83d.

82. Jing Li et al., "Advances in Synthetic Biology and Biosafety Governance," *Frontiers in Bioengineering and Biotechnology* 9 (April 2021): 598087, https://doi.org/10.3389/fbioe.2021.598087.

83. Erik Parens, Josephine Johnston, and Jacob Moses, "Ethical Issues in Synthetic Biology: An Overview of the Debates," Woodrow Wilson International Center for Scholars, June 2009, https://www.wilsoncenter.org/sites/default/files/media/documents/publication/synbio3.pdf.

84. Drew Endy and Laurie Zoloth, "Should We Synthesize a Human Genome?," *Dspace@MIT*, May 10, 2016, https://dspace.mit.edu/handle/1721.1/102449.

85. "Home," National Security Commission on Emerging Biotechnology, accessed September 3, 2024, https://www.biotech.senate.gov.

86. Drew Endy and Mike Kuiken, "US Must Embrace a Winning Biotech Strategy," *Boston Globe*, February 21, 2025, https://www.bostonglobe.com/2025/02/21/opinion/us-china-biotechnology-innovation-manufacturing/.

87. Iris Luo, "China's Biopharma Boom in Global Drug Licensing Deals," EC Innovations, July 3, 2025, https://www.ecinnovations.com/blog/chinas-biopharma-boom-in-global-drug-licensing-deals/.

STANFORD EXPERT CONTRIBUTORS

Dr. Drew Endy
SETR Faculty Council and Martin Family Faculty Fellow in Undergraduate Education (Bioengineering)

Dr. Dan Lin-Arlow
Cofounder and Chief Scientific Officer at Ansa Biotechnologies, Inc.

Dr. Possu Huang
Assistant Professor of Bioengineering

Dr. Brian Hie
Assistant Professor of Chemical Engineering

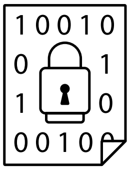
Dr. Michael Jewett
Professor of Bioengineering

Dr. Jennifer Brophy
Assistant Professor of Bioengineering

Samuel Hyo-nam King
PhD Student in Bioengineering

Annie Nguyen
SETR Fellow and PhD Student in Bioengineering

03



CRYPTOGRAPHY AND COMPUTER SECURITY

KEY TAKEAWAYS

- Cryptography is essential for protecting information, but alone it cannot secure cyberspace against all threats; it must operate in concert with the broader field of computer security.
- Cryptography is the enabling technology of blockchain, which is the enabling technology of cryptocurrencies.
- Rather than pursue a central bank digital currency, the United States has adopted a policy preference for privately issued digital assets, promoting stablecoins and cryptocurrencies as vehicles for financial innovation and resilience.

Overview

The word *cryptography* originates from Greek words that mean "secret writing." Once limited to simple codes and ciphers, it now relies on advanced mathematics to protect data from unauthorized access or tampering.[1] Though largely invisible, cryptography secures many everyday interactions, from online shopping to cell phone calls.

Cryptography is essential for internet activity—from messaging and banking to everyday browsing—but it cannot, on its own, guarantee the confidentiality, integrity, or availability of information. Various vulnerabilities ensure that cybersecurity will remain an ongoing challenge. These include technical vulnerabilities in the digital systems that humans operate and use; human vulnerabilities, such as the tendency to bypass security mechanisms because using them is considered inconvenient; and strong incentives for attackers.

Cryptography is essential for internet activity—from messaging and banking to everyday browsing.

Cryptography Basics: Public Keys, Private Keys, and Hashes

Here's an example of how cryptography works: Drew wants to send a private message to Taylor. She scrambles (encrypts) the message using an encryption algorithm and sends the scrambled (ciphertext) version. When Taylor receives it, he unscrambles (decrypts) it to recover the original (plaintext) message. Ellen, an eavesdropper, tries to intercept the message and must find a way to break the cryptographic protection to see the plaintext.

One example of an encryption algorithm is the shift cipher, where each letter is replaced by one *N* positions later in the alphabet. If $N = 2$, A in the plaintext becomes C in the ciphertext, B becomes D, and so on. If $N = 3$, A becomes D. To decrypt the ciphertext, Taylor must know that the algorithm was the shift cipher and the key *N*—so if he sees C and knows $N = 2$, he writes down A. (Modern encryption is far more secure and complex than this example but also harder to explain.)

Both Drew and Taylor must share a secret: *N*, the cryptographic key—a string of digits used for both encryption and decryption. They must also know that the algorithm is the shift cipher. If Ellen learns both the algorithm and key, she can decrypt the message. This type of encryption, where the same key is used by both parties, is known as symmetric or secret-key cryptography. It requires secure key distribution—a way to share keys with intended recipients while keeping them from others.

Symmetric cryptography poses a practical problem: Parties have to meet in person to exchange secret keys before communicating securely. Imagine having to meet every phone contact face to face before speaking. In the 1970s, Stanford professor Martin Hellman and Whitfield Diffie introduced asymmetric, or public-key, cryptography. This uses two keys: a public key, which anyone can use to encrypt a message and can be distributed over insecure channels, and a private key, known only to the intended recipient (see figure 3.1), which is needed to decrypt it. Although the keys are mathematically linked, deriving the private key from the public key would take longer than the age of the universe (unless quantum computing changes that, as discussed later in this chapter; for an in-depth discussion of quantum computing, see chapter 7, on quantum technologies).

Cryptography also enables the creation of secure hashes. A hash accepts a message of any length and computes a unique fixed-length string of numbers—called the hash value—corresponding to that message. Hashes have two key properties: It is extremely difficult to find another message with the same hash value, and it is infeasible to recover the original message from the hash value alone.

Using a secure hash function, the sender can use public-key cryptography to ensure integrity (protection against tampering) and identity (the message originated from the stated sender).

To illustrate, Alice (the sender) first computes the hash value of her message. Next, she encrypts the hash value with her private key, a process analogous to signing a document, generating a digital signature of the message's hash.[2] Alice then sends the message and its digital signature to Bob (the receiver).

Once Bob receives it, he can recover the hash value for the message that Alice purportedly sent and compare that value to his own computation of the hash value. If these match, Bob can be assured that the message has not been altered in transmission and that Alice sent it, since only Alice could have used her private key to create a digital signature of the message's hash.

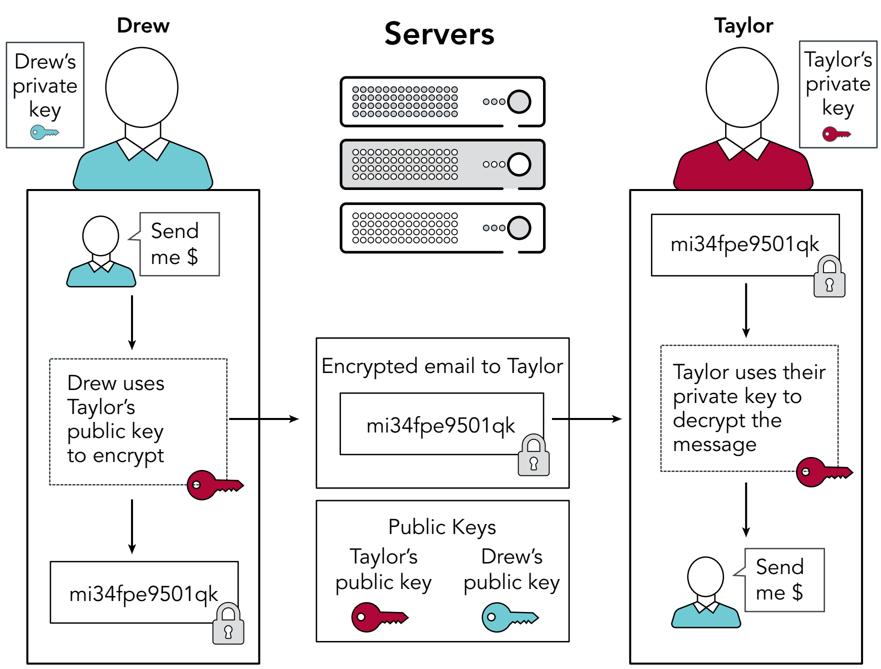
Messages can also be digitally time stamped. A known authoritative time and date server—such as the Internet Time Service, operated by the National Institute of Standards and Technology—accepts a message, appends the current date and time, and then provides a digital signature for the stamped message.

Computer Security

Computer security traditionally focuses on safeguarding computer systems against unauthorized access and misuse. It emphasizes the core principles of confidentiality, integrity, and availability—with all three also known collectively as the CIA triad. Confidentiality refers to the privacy of data (i.e., preventing unauthorized disclosure). Integrity refers to preserving data (i.e., guarding against unauthorized alterations). Availability refers to data and resources being accessible to authorized users, especially during critical times.

Historically, computer security focused on protecting individual machines from actions perpetrated by malicious actors, whether individuals or states. Over time, the focus has expanded—first to securing the infrastructure of increasingly networked systems and now to addressing vulnerabilities in machine learning (ML) models. Cryptography is one of many tools whose use can enhance computer security. However, the protections afforded by even perfect cryptography can often be circumvented by taking advantage of vulnerabilities in the computer systems on which

FIGURE 3.1 How public-key cryptography works



that cryptography is implemented. For example, if an intercepted encrypted message is too hard to decrypt, the attacker's focus will most likely be on exploiting vulnerabilities in computer security to obtain the message before the sender encrypts it or after the receiver decrypts it.

Thus, cryptography and security are inseparable, and using cryptography is by no means a guarantee of security.

Blockchain Technology

A blockchain is a chain of digital blocks, each containing a transaction and a cryptographic hash of the previous block. This links every block (except the first) to its predecessor. As new transactions occur, new blocks are added, extending the chain.

Blockchains are distributed across thousands of computers, ensuring they are highly decentralized. They enable multiple parties to coordinate transactions without a central trusted authority—a common need in financial settings. Transactions recorded on them cannot be altered retroactively without detection. Because blockchains are widely distributed across thousands of computers, they are always accessible: Anyone can deploy or interact with applications, and no one can block access to them. Data on blockchains cannot be erased; later transactions may correct errors, but the original record remains.

The distributed nature of blockchains also increases security. A new transaction on a blockchain is broadcast to every party in the network, each of which has a replica of the entire blockchain (see figure 3.2). Each party then tries to validate the new transaction. These replicas may not be fully synchronized; some might have received the new transaction, while others may have not. To ensure that all replicas are identical, blockchains use consensus mechanisms to agree on the correct information. Ethereum, for example, accepts transactions that have been validated by two-thirds of the participants. Blockchains are designed with economic incentives for replicas to behave honestly.

Applications that run on a blockchain are called smart contracts—computer programs that are always available and whose execution cannot be reversed. They can implement financial instruments, record ownership of digital assets, or support marketplaces for buying and selling. Smart contracts are also composable: One contract can use another, enabling a vibrant ecosystem where projects build on each other. Once deployed, the contracts remain available indefinitely. This is in contrast to cloud applications, which disappear when developers stop paying hosting fees.

Key Developments

A Host of Blockchain Applications

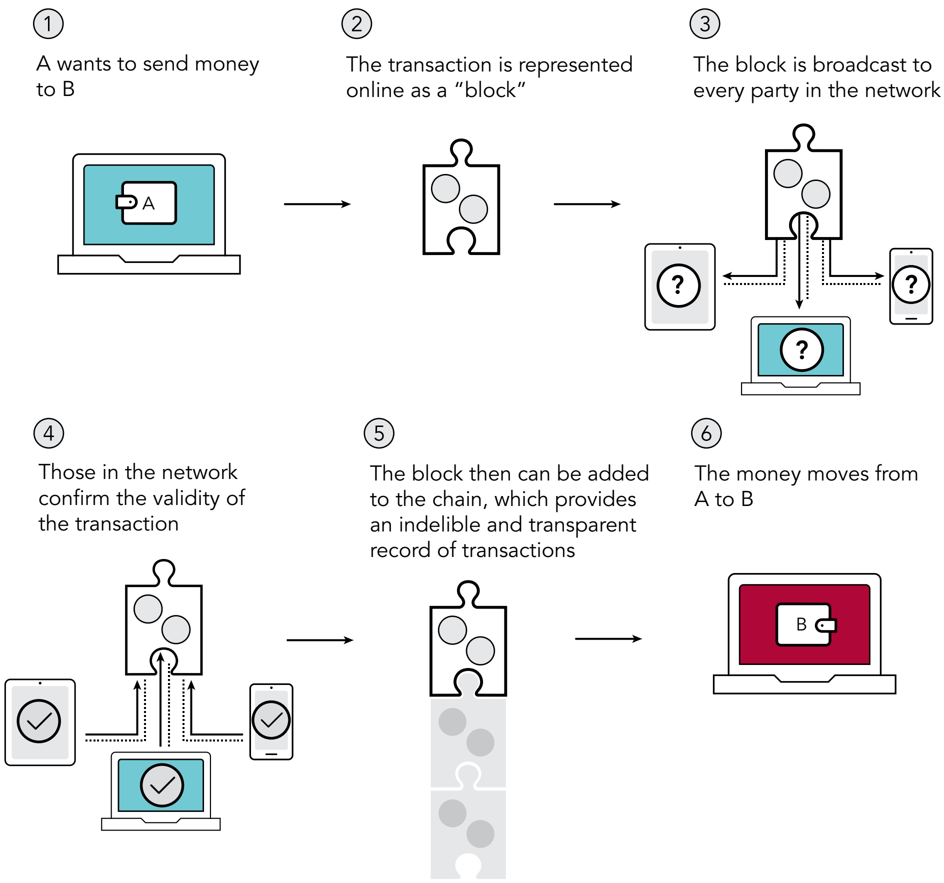
Blockchain technology was developed decades ago but has recently been used for a variety of applications. Many of these are operational today, though they are often at limited scale. (For a more comprehensive discussion of examples, see the chapter on cryptography in the 2025 *Stanford Emerging Technology Review* [*SETR*].) Some current examples include the following:

- **Time stamping and data provenance** Because data written to a blockchain cannot be modified or removed, blockchains provide a secure mechanism for data provenance and time stamping. For instance, creators can post cryptographic hashes of work to a blockchain to establish authorship or creation dates.

- **Identity management** Blockchains enable secure storage and selective disclosure of personal records (e.g., diplomas, birth certificates, financial records), allowing users to prove facts—such as being of a certain age—without revealing sensitive details (such as their actual age). One such application, SpruceID, is already being deployed.[3]

- **Supply chain management** Blockchains provide a transparent and secure way to track the

FIGURE 3.2 How a blockchain manages transactions



movement of goods and their origin and quantity in industries ranging from luxury goods to food labeling.

- **Transactional records** Storing contracts or sales records on blockchains can reduce fraud, simplify auditing, and streamline operations.

- **Cryptocurrencies** These are digital instruments that many people use as a medium of exchange. Well-known ones include Bitcoin, Ethereum, Avalanche, and Polygon, each of which has its own unique features and applications. Because they are not issued by any central authority, they are not subject to the same national regulatory regimes that govern traditional currencies (i.e., so-called fiat currencies).

Cryptocurrencies use a blockchain structure to ensure the integrity and immutability of transaction data, making it resistant to fraud and counterfeiting and reducing its susceptibility to government interference or manipulation. Contrary to a common belief, cryptocurrencies can, but do not have to, support private or secret transactions. Indeed, the most popular cryptocurrencies deliberately do not hide the details of their transactions. Those who transact in cryptocurrencies often wish to exchange their instruments for fiat currency (e.g., real dollars) and generally use a cryptocurrency exchange to do so.

Secure Computation

The field of cryptography has also expanded in scope to include secure computation, which enables multiple parties to jointly compute functions where the inputs from each party are kept secret from the others. Secure computation enables data privacy during computation, ensuring that no party learns more information about the other parties' inputs than what can be inferred from the result alone. It also allows users to prove they possess knowledge of a statement without having to disclose the actual content of that statement. (For a more detailed explanation of secure computation with illustrative examples and explanations, see the chapter on cryptography in *SETR* 2025.) A few representative applications include the following:

- **Private statistics** Stanford's Prio system lets users contribute data, such as COVID-19 exposure status, to an aggregate total without disclosing individual responses.[6]

- **Financial privacy** Banks can collaborate to detect fraud patterns across institutions without revealing individual customer records.

- **Privacy-preserving auctions** These can determine a winner without exposing losing bids, maintaining fairness while protecting private financial information.

Zero-Knowledge Proofs

Zero-knowledge proofs are cryptographic protocols that allow one party (the prover) to convince another (the verifier) that a statement is true without revealing why it is true. For example, someone can prove they know a password or have enough funds for a purchase without disclosing the password or the amount of money. This privacy-preserving technique has moved from theory into real-world applications, such as the following:

- **Banking** The cryptocurrency Zcash uses zero-knowledge proofs to let users prove they can afford a transaction without having to reveal their account balance.[4]

- **Provenance for digital images** The Coalition for Content Provenance and Authenticity employs zero-knowledge proofs to ensure that an image was captured by a verified camera and underwent only permitted edits—without trusting the editing software itself.[5]

- **Cooperative tracking and verification of numbers of tactical nuclear warheads** Experimental systems have used zero-knowledge proofs to track changes in warhead status while concealing sensitive military information. Though the use has not yet been adopted in formal treaties, its feasibility in principle has been demonstrated.[6]

A more detailed introduction to zero-knowledge proofs and their use cases is available in the chapter on cryptography in *SETR* 2025.

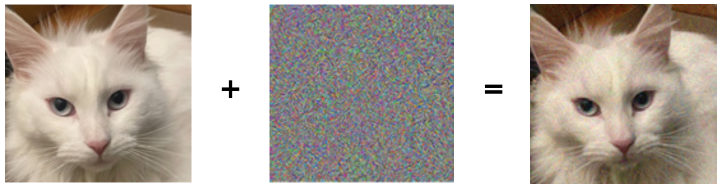
Over the Horizon

Impact of Cryptography

The applications described above suggest a broad range of possibilities for cryptographically enabled data management services. Whether we will see their widespread deployment depends on complicated decisions about economic feasibility, costs, regulations, and ease of use.

Misaligned incentives can affect how fast innovations are deployed. Some of the applications described above provide significant benefits for the parties whose data can be better protected and kept more private. But existing companies, having built their

FIGURE 3.3 How adversarial noise can fool image classifiers



Cat + Adversarial noise = Guacamole

Adding carefully crafted noise (center) to an image of a cat (left) produces an altered version (right) that looks identical to humans but can cause a model to misclassify it—for example, as guacamole.

Source: Neil Alexander Perry

business models on legacy systems that ingest all their customers' data, have no incentive to change their practices. They are the ones who would have to pay for these privacy-protecting capabilities, yet they would not benefit from their adoption.

Widespread deployment will also require confidence that proposed innovations will work as advertised (i.e., would-be users of these innovations must have confidence in them). But concepts such as secure computation and zero-knowledge proofs are math heavy and counterintuitive to most people. Getting policymakers, consumers, and regulators to place their trust in these applications will be challenging.

Machine Learning Security: Adversarial Risks and Systemic Vulnerabilities

As ML systems move into high-stakes settings—including autonomous vehicles, financial platforms, and healthcare diagnostics—their security under adversarial conditions is becoming a critical concern (for more, see chapter 1, on AI). In an ML system, small, malicious changes to inputs can cause large, unexpected model failures. These brittle responses undermine trust in a system's ability to operate safely in environments where reliability is paramount.

This fragility stems from a core asymmetry. While ML performs well on the inputs that most average users would give it, it often fails on inputs that are crafted by deliberately malicious adversaries. In other words: ML systems are great for random data, but they often perform poorly when confronted with deliberately crafted adversarial data.

ATTACKS ON ML SYSTEMS

Researchers have identified attacks targeting every stage of the ML pipeline:

- Training-time attacks corrupt models during learning. Carefully altering even a single image in a dataset of ten thousand can lead to persistent misclassifications—such as labeling images of dogs as fish.[7] These techniques, once confined to research, now appear in the real world. The tool Nightshade deliberately allows artists to corrupt images before posting them online, sabotaging unauthorized AI training on scraped content.[8]

- Inference-time attacks occur after deployment. An attacker may introduce "noise"—tiny, imperceptible modifications to data that cause the model to produce incorrect outputs.[9] For example, making small alterations to the pixels in a cat image can make the model label the image as guacamole (figure 3.3). The key point is that such attacks—repeatedly shown to be possible over the past several years—demonstrate that a model can sometimes be tricked into errors that would be obvious to humans. Imagine the risk if an adversary could make a military reconnaissance system mistake tanks for school buses or induce an airport security scanner to mistake a gun for a notebook.

These threats extend beyond images. For example, because large language models (LLMs) cannot distinguish between inputs intended as data versus inputs intended as commands, they may interpret a phrase embedded in data as a command, which may be hostile or malicious rather than benign. For example, this strategy—which is an instance of a well-known hacker technique known as prompt injection—could result in getting LLMs to leak confidential information, ignore safety constraints, or perform unintended actions.[10] As these models become linked to tools like email and payments, such attacks carry more risk.[11] As AI agents—autonomous AI software programs that have access to important data or controls—become increasingly popular, there is a growing risk that these could be "tricked" by malicious content on the internet.

EMERGING RISKS AND THE SECURITY GAP

The rapid deployment of AI systems has outpaced available security solutions. Traditional techniques from computer security and common defenses like input filtering offer only partial protection and often shift the vulnerability elsewhere. For example, if inputs are digitally filtered for safety, attackers may instead target the software filters themselves, which are often susceptible to similar exploits. Some researchers are now exploring new defenses for inference-time threats, such as isolation to protect sensitive components and data from being compromised by malicious or untrusted inputs.[12] Another defense involves using stricter control flows that explicitly manage how decisions, loops, branches, and data interactions occur to help ensure a system's predictable, secure, and reliable operation.[13]

To secure the training process, other efforts focus on hardware safeguards, such as trusted execution environments—secure zones within a host system that preserve data confidentiality and computational integrity even if the system is compromised. Stanford researchers are developing auditable training pipelines that log each intermediate step in training. This enables users to verify the model's training process (e.g., to ensure that the data on which it was trained was not compromised in some way) and trace certain security issues back to their origin when problems arise.[14]

All of these defenses remain in their early stages. No current approach offers broad protection across all tasks, data types, or adversarial techniques. The field remains in an arms race: New attacks emerge rapidly, while robust, scalable defenses continue to lag behind. In this landscape, any claim to deploy ML to solve a problem should prompt an immediate question: What have you done about adversarial inputs and attacks?

DUAL-USE CAPABILITIES AND MODEL INTEGRITY

LLMs raise classic dual-use concerns. Their ability to identify software vulnerabilities can assist defenders in fixing systems—or can arm attackers to more easily exploit vulnerabilities. Studies show LLM-based agents can already solve many standardized cybersecurity tasks, rivaling novice human hackers.[15] Whether they will ultimately favor offense or defense remains uncertain. Their use in software development can accelerate productivity but also create new vulnerabilities. For example, LLMs often generate insecure or outdated code, especially when they are used by nonexperts who lack awareness of best practices.[16]

Even after they are deployed, models remain vulnerable to extraction attacks. These involve adversaries reconstructing similar models through repeated queries of a target model, enabling them to gather training data that the original model uses.[17] This threatens both intellectual property and the safeguards meant to prevent misuse of the target model. This issue is further compounded

Bitcoin mining uses more energy than the Netherlands.

by the phenomenon of transferability, where an attack on one model often works on similar models. Transferability means that attackers don't need internal access to the original model to succeed. It also means that similar models can be constructed to aid in the development of attacks on the original model, regardless of the protections and safeguards embedded in the original.

Policy Issues

Research Infrastructure

Although cryptography is fundamentally a mathematical discipline, it requires both human talent and substantial computing resources to examine the efficiency of new techniques, write computationally expensive software such as zero-knowledge provers, and conduct comprehensive scans of the internet. Progress also relies on interdisciplinary centers that bring together faculty from different fields to share problem sets and understand the potential benefits of cryptographically enabled techniques and approaches.

Research is funded by both the US government and private industry, but funding from the US government is subject to many requirements that increase the difficulty of proposal submission manyfold (as much as by a factor of sixty). Thus, research faculty often prefer arrangements with the private sector, which tend to be much simpler. On the other hand, only the US government is able to fund research that may not pay off for many years (as in the case of quantum computing).

EXCEPTIONAL ACCESS

Exceptional access regulations would require communications carriers and technology vendors to provide US law enforcement agencies access to encrypted information (both data storage and communications) under specific legal conditions. Opponents of exceptional access argue that implementing this capability inevitably weakens the security afforded by encryption to everyone. Supporters of exceptional access do not debate this technical assessment: It is true that exceptional access, by definition, weakens encryption. However, they argue that even if lower security is the result of implementing exceptional access, that price is worth the benefits to law enforcement.[18]

ENERGY CONSUMPTION

Bitcoin, an older cryptocurrency and today the dominant one, consumes an enormous amount of energy; Bitcoin mining uses more energy than the Netherlands.[19] For this reason, newer blockchains—notably Ethereum—are designed to use far less energy; today Ethereum's annual energy use is less than a ten-thousandth of YouTube's annual consumption. But Ethereum's market capitalization is less than half that of Bitcoin, and it remains to be seen whether any less energy-intensive cryptocurrency will displace the latter.

QUANTUM COMPUTING AND CRYPTOGRAPHY

Current public-key cryptography is based on the extraordinarily long times—ones comparable to the age of the universe—today's computers require to derive a private key from its public-key counterpart. When realized, quantum computing (discussed more fully in chapter 7, on quantum technologies) will pose a significant threat to today's public-key algorithms. Experts disagree on how long it will take to build quantum computers that are capable of this, but under the May 2022 National Security Memorandum 10, Promoting US Leadership in Quantum Computing While Mitigating Risks to Vulnerable Cryptographic Systems, the US government has initiated the transition to quantum-resistant public-key algorithms. Many experts in the field expect quantum-resistant algorithms will be widely available by the time quantum computing comes online.

At the intersection of quantum computing and cryptography are two important issues: (1) that support for the transition to a quantum-resistant encryption environment should continue with urgency and focus, and (2) that messages protected by pre-quantum cryptography will be vulnerable in a post-quantum world. If those messages have been saved by adversaries (which is likely in the case of parties like Russia), those bad actors will be able to read a host of old messages. Containing secrets from the past, they may reveal embarrassments and dangers with potentially detrimental policy implications.[20]

CRYPTOCURRENCIES AND THE EMERGING US POLICY APPROACH

While many countries are pursuing central bank digital currencies (CBDCs) to modernize their financial systems, the United States is taking a markedly different path. CBDCs are cryptography-based currencies issued by central banks, with legal tender status and value tied to a nation's traditional currency. They promise fast, low-cost payments with centralized oversight. Advocates cite benefits such as greater financial inclusion, lower cross-border costs, and preserving the dollar's global role—especially as rivals like China advance their own CBDCs. Critics, however, warn of privacy risks, centralized surveillance, and excessive government control. (For a full discussion of CBDCs, see the chapter on cryptography in *SETR* 2025.)

Departing from earlier policy, the Trump administration has signaled support for privately issued cryptocurrencies over a CBDC. In January 2025, it issued an Executive Order (EO) on Digital Assets that revoked President Biden's 2022 EO 14067, which had promoted CBDC exploration, consumer protection, and anti–illicit finance measures. The new order explicitly prohibits the development or promotion of a US CBDC and establishes an interagency working group to coordinate digital asset policy and regulation.

These policy shifts have occurred amid ongoing debate over how to classify and regulate digital assets. The 2023 collapse of cryptocurrency exchange FTX and the subsequent conviction of its founder Sam Bankman-Fried intensified scrutiny over whether cryptocurrencies should be treated as securities or currency. Despite some legislative progress, many investors, consumers, and entrepreneurs remain uncertain about their regulatory status.

A March 2025 EO established a US Strategic Bitcoin Reserve and digital asset stockpile, using Bitcoin seized from illegal activities and directing agencies to pursue additional budget-neutral acquisitions of the cryptocurrency. Supporters see it as a way to diversify national reserves, hedge against inflation, and promote US leadership in digital asset innovation. Critics point to Bitcoin's price volatility, its limited utility in crises and cybersecurity risks, and the risk of potential conflicts of interest that could undermine public trust, given that the policymakers themselves may have significant cryptocurrency holdings.

In addition to the above activity, the Guiding and Establishing National Innovation for US Stablecoins (GENIUS) Act, signed in July 2025, created a federal framework for issuing payment stablecoins—cryptocurrencies that are designed to have stable prices. By combining the speed and programmability of cryptocurrencies with the familiarity of fiat currency (i.e., ordinary money), stablecoins allow users to transact without the price volatility typically associated with many cryptocurrencies and other digital assets. However, the law requires stablecoins to be backed by reserve assets such as Treasury bills and precious metals rather than algorithmic mechanisms to enhance their price stability. This may limit innovation via experimentation with alternative cryptocurrency designs.

The Trump administration is also backing the Digital Asset Market Clarity Act,[21] or the Clarity Act, which proposes making the Commodity Futures Trading Commission the primary regulator of digital commodities and their intermediaries while maintaining certain Securities and Exchange Commission (SEC) powers over initial crypto sales. The act would also introduce a special exemption that eases SEC registration requirements for fundraising purposes. Under this proposed legislation,[22] digital commodities would be defined as digital assets whose worth is "intrinsically linked" to their activity on a blockchain. This group includes nearly all cryptocurrencies in use today. However, the definition of the term "digital commodity" would exclude securities, derivatives, and stablecoins, even if they are based on blockchains.

Against this backdrop, cryptocurrencies pose complex and evolving policy challenges. These include the following:

- **Regulatory clarity and market integrity** Regulatory ambiguity and weak oversight continue to be a challenge facing digital asset markets. The decentralized, cross-border nature of cryptocurrencies complicates efforts to classify and supervise them. Inherent volatility of cryptocurrencies, combined with limited transparency across many exchanges, exposes users to fraud, manipulation, and financial risk. As use of these assets grows, so do calls for clearer rules, better disclosures, and stronger consumer protections. Tax reporting remains another challenge: The pseudonymous nature of cryptocurrencies complicates enforcement, and users often lack a clear understanding of their tax-reporting obligations. The GENIUS and Clarity Acts are first steps toward regulatory clarity, but as the use of cryptocurrencies expands, the surfacing of other issues requiring further legislative and executive branch attention is inevitable.

- **Financial crime and illicit activities** The pseudonymous nature of cryptocurrencies and their cross-border use also enables or facilitates money laundering, tax evasion, and sanctions evasion, creating major enforcement challenges. Authorities are expanding international cooperation, tightening anti–money laundering and know-your-customer protocols related to cryptocurrencies, and working to close regulatory gaps.

- **Economic and monetary policy risks** Because cryptocurrencies bypass traditional financial systems, their widespread use weakens central banks' ability to control the money supply and set interest rates across the economy. If cryptocurrencies are integrated into mainstream finance (e.g., through retirement funds, banking systems, or national reserves), a collapse in cryptocurrency valuations could trigger a financial crisis, impacting savings and investments across the economy. Additionally, as more people use cryptocurrencies instead of fiat money, confidence in government-issued currencies may erode.

- **Conflicts of interest and governance transparency** Government actions can significantly affect cryptocurrency prices (as is true of any other investment asset), raising concerns about personal financial gain among policymakers and regulators. Industry influence over the regulatory process also prompts political and ethical scrutiny.

- **Cybersecurity risks** The decentralized architecture of cryptocurrencies creates novel opportunities for cyberattacks across digital wallets, crypto exchanges, and smart contracts. Hacks, phishing, and other exploits may cause substantial losses and erode trust. As crypto assets intertwine with traditional finance, their vulnerabilities may trigger broader economic fallout.

Strong cybersecurity standards, incident reporting, and federal coordination are essential to limit systemic risk to the national and global financial system.

- **Privacy and surveillance concerns** Digital asset regulation increasingly intersects with debates over financial privacy and civil liberties. Expanding anti–money laundering and know-your-customer rules may lead to calls for digital identity systems, raising concerns about surveillance and state overreach. At the same time, technologies that enhance privacy may face increased scrutiny. Policymakers will need to carefully balance law enforcement needs with privacy concerns.

NOTES

1. "Cryptography," National Institute of Standards and Technology, US Department of Commerce, n.d., https://www.nist.gov/cryptography.

2. In this context, encrypting the hash value simply means running the encryption algorithm using a string of numbers that just happen to be Alice's private key as the input. In most cases involving public-key cryptography, the private key is used only for decryption purposes, but nothing stops a user from using it in other ways.

3. "SpruceID," Spruce Systems, accessed October 13, 2024, https://spruceid.com/.

4. "What Are Zero-Knowledge Proofs?," Zcash, accessed September 7, 2025, https://z.cash/learn/what-are-zero-knowledge-proofs.

5. Trisha Datta and Dan Boneh, "Using ZK Proofs to Fight Disinformation," Medium, September 29, 2022, https://medium.com/@boneh/using-zk-proofs-to-fight-disinformation-17e7d57fe52f.

6. Miles A. Pomper, William Alberque, Marshall L. Brown Jr., William M. Moon, and Nikolai Sokov, "Everything Counts: Building a Control Regime for Nonstrategic Nuclear Warheads in Europe," 2022, CNS Occasional Paper #55, Monterey, CA: James Martin Center for Nonproliferation Studies, https://nonproliferation.org/op55-everything-counts-building-a-control-regime-for-nonstrategic-nuclear-warheads-in-europe.

7. Pang Wei Koh and Percy Liang, "Understanding Black-Box Predictions via Influence Functions," *Proceedings of the 34th International Conference on Machine Learning*, in *Proceedings of Machine Learning Research* 70 (2017): 1885–94, https://proceedings.mlr.press/v70/koh17a.html.

8. Shawn Shan, Wenxin Ding, Josephine Passananti, Stanley Wu, Haitao Zheng, and Ben Y. Zhao, "Nightshade: Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models," *2024 IEEE Symposium on Security and Privacy*, 2024, 807–25, https://doi.org/10.1109/SP54263.2024.00207.

9. Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy, "Explaining and Harnessing Adversarial Examples," preprint, arXiv, December 20, 2014, https://arxiv.org/abs/1412.6572; Anish Athalye, Logan Engstrom, Andrew Ilyas, and Kevin Kwok, "Synthesizing Robust Adversarial Examples," *Proceedings of the 35th International Conference on Machine Learning*, 2018, 284–93, https://proceedings.mlr.press/v80/athalye18b.html.

10. Hezekiah J. Branch, Jonathan Rodriguez Cefalu, Jeremy McHugh, Leyla Hujer, Aditya Bahl, Daniel del Castillo Iglesias, Ron Heichman, and Ramesh Darwishi, "Evaluating the Susceptibility of Pre-Trained Language Models via Handcrafted Adversarial Examples," preprint, arXiv, September 5, 2022, https://arxiv.org/abs/2209.02128.

11. Kai Greshake, Sahar Abdelnabi, Shailesh Mishra, Christoph Endres, Thorsten Holz, and Mario Fritz, "Not What You've Signed Up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection," *Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security*, November 26, 2023, 79–90.

12. Luca Beurer-Kellner, Beat Buesser, Ana-Maria Creţu, Edoardo Debenedetti, Daniel Dobos, Daniel Fabian, Marc Fischer, et al., "Design Patterns for Securing LLM Agents Against Prompt Injections," preprint, arXiv, last updated June 27, 2025, htt ps://arxiv.org/abs/2506.08837.

13. Edoardo Debenedetti, Ilia Shumailov, Tianqi Fan, Jamie Hayes, Nicholas Carlini, Daniel Fabian, Christoph Kern, Chongyang Shi, Andreas Terzis, and Florian Tramèr, "Defeating Prompt Injections by Design," preprint, arXiv, last updated June 24, 2025, https://arxiv.org/abs/2503.18813.

14. Megha Srivastava, Simran Arora, and Dan Boneh, "Optimistic Verifiable Training by Controlling Hardware Nondeterminism," *Proceedings of the 38th International Conference on Neural Information Processing Systems*, 2024 (Curran Associates, Inc., 2025), 95639–61, https://dl.acm.org/doi/10.5555/3737916.3740946.

15. Andy K. Zhang, Neil Perry, Riya Dulepet, Joey Ji, Celeste Menders, Justin W. Lin, Eliot Jones, et al., "Cybench: A Framework for Evaluating Cybersecurity Capabilities and Risks of Language Models," preprint, arXiv, 2024, https://arxiv.org/abs/2408.08926.

16. Neil Perry, Megha Srivastava, Deepak Kumar, and Dan Boneh, "Do Users Write More Insecure Code with AI Assistants?," *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, 2023 (Association for Computing Machinery, 2023), 2785–99, https://doi.org/10.1145/3576915.3623157.

17. Florian Tramèr, Fan Zhang, Ari Juels, Michael K. Reiter, and Thomas Ristenpart, "Stealing Machine Learning Models via Prediction APIs," *25th USENIX Security Symposium*, 2016, preprint, arXiv, 2016, 601–18, https://arxiv.org/abs/1609.02943.

18. "Attorney General William P. Barr Delivers Keynote Address at the International Conference on Cybersecurity," Office of Public Affairs, US Department of Justice, July 23, 2019, https://www.justice.gov/opa/speech/attorney-general-william-p-barr-delivers-keynote-address-international-conference-cyber.

19. "Bitcoin Energy Consumption Index," Digiconomist, accessed September 7, 2025, https://digiconomist.net/bitcoin-energy-consumption.

20. Herbert Lin, "A Retrospective Post-Quantum Policy Problem," Lawfare, 2022, https://www.lawfaremedia.org/article/retrospective-post-quantum-policy-problem.

21. Although the vast majority of cryptocurrencies are indeed based on blockchain, it is possible in principle to build a cryptocurrency on a different underlying technology.

22. Paul Tierno, Crypto Legislation: An Overview of H.R. 3633, the CLARITY Act, Congressional Research Service, Library of Congress, July 7, 2025, https://www.congress.gov/crs-product/IN12583.

STANFORD EXPERT CONTRIBUTORS

Dr. Dan Boneh
SETR Faculty Council and Professor of Computer Science and of Electrical Engineering

Dr. David Tse
Thomas Kailath and Guanghan Xu Professor of Engineering

Neil Perry
SETR Fellow and PhD Student in Computer Science

04

ENERGY TECHNOLOGIES

KEY TAKEAWAYS

- Although many clean energy technologies are now available and increasingly affordable, scaling them up and building the infrastructure for them will take decades due to infrastructure inertia, stakeholder complexity, and the "energy trilemma," which balances reliability, affordability, and cleanliness.

- The US has shifted from climate urgency to energy dominance, redirecting support from renewables and electric vehicles to fission, coal, and natural gas. Globally, similar trends prevail as nations record peak fossil fuel use and scale back renewable investments, prioritizing energy security over decarbonization.

- Energy innovation is fragmented, diverse, and geopolitically strategic, with progress in technologies like fission, geothermal, fusion, and batteries reshaping the energy frontier. To compete with China, US technology leadership depends on sustained research and development funding, robust supply chains, and strategic industrial policies.

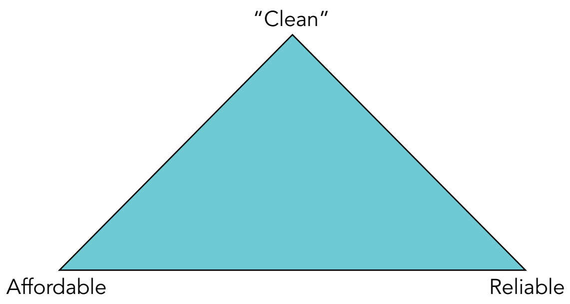
Overview

Energy is the lifeblood of modern society—enabling heating, cooling, light, mobility, information, and the creation of modern materials. Because it touches everything, everywhere, all the time, energy plays out against a complex backdrop of technology, economics, regulation, and consumer behavior. Key elements of this backdrop include the following:

- **Growing demand** As several billion people in the developing world lift themselves out of poverty, global energy consumption is projected to increase by some 50 percent between 2020 and 2050.[1] That increase is not a luxury but is essential to their improved quality of life.

- **The "energy trilemma"** It's not enough that energy systems produce and deliver energy. They need to do so reliably, affordably, and cleanly, with "clean" referring to both local and greenhouse gas

FIGURE 4.1 The energy trilemma



emissions. (Local emissions refer to particulates emitted in the immediate vicinity of a power plant.) Those three dimensions are often expressed as the energy trilemma, as illustrated in figure 4.1.

It is rare to find technologies that simultaneously satisfy all three desiderata. In the US electricity sector, conventional coal is secure and affordable but generally emits greenhouse gases; natural gas is much cleaner locally but still emits carbon dioxide (CO_2); wind and solar are affordable and non-emitting but unreliable; and nuclear power is both clean and reliable but more expensive than alternatives. The trilemma suggests that no single type of energy source will always be right under all circumstances.

- **Opportunities for innovation** The challenge of resolving the energy trilemma has engendered a flurry of technological innovation. That effort has dramatically reduced the costs of onshore wind and solar generation, improved battery performance and economics, surfaced promising geothermal technologies, and rekindled interest in nuclear power, particularly designs for small reactors. Although the many innovations on today's drawing boards will not be impactful for years, they are the foundation for a more affordable, more reliable, and cleaner energy future in the longer term.

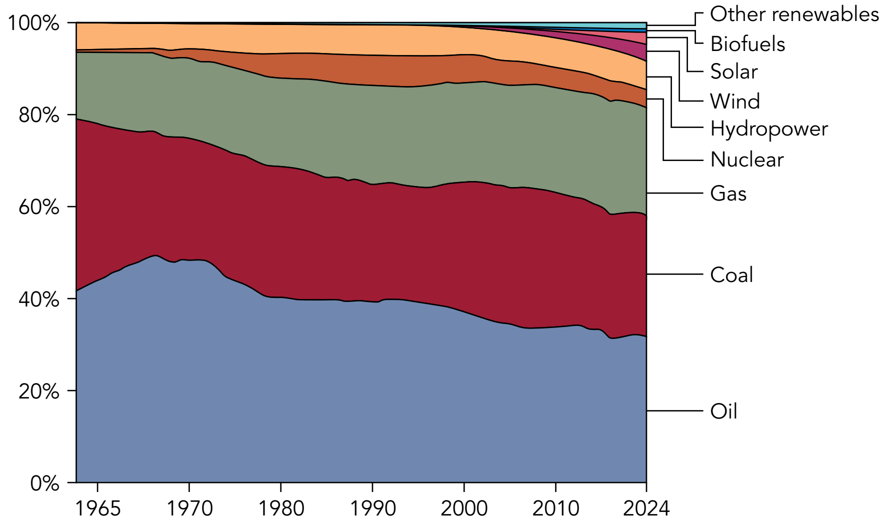
- **Growing electrification** Energy is delivered to end users through carriers, whether they be fuel molecules or electrons. The latter carrier is favored in more advanced countries. This is because electricity is easily moved through wires, is clean at the point of use, and can be employed in many ways, from powering electronics to moving electric vehicles (EVs). Global electricity demand grew 4.3 percent in 2024, almost double the average rate over the prior decade. The rise of data centers, heat pumps, and vehicle electrification is expected to increase US electricity demand by some 25 percent in the next five years, a dramatic acceleration compared with the past two decades, when demand was essentially flat.[2]

- **Hydrocarbon dominance** As figure 4.2 shows, hydrocarbons derived from fossil fuels (coal, oil, and natural gas) supplied 86 percent of the world's primary energy in 2024.[3] (Primary energy refers to energy sources before they have been converted to electricity.) Wind and solar generation, while growing rapidly, accounted for 6.5 percent of primary energy the same year.

- **Limitations of renewable sources** Wind- and solar-generated electricity remain substantially cheaper than electricity generated by fossil fuels and also accounted for more electrical energy in 2024 than in any previous year.[4] However, the

The challenge of resolving the energy trilemma has engendered a flurry of technological innovation.

FIGURE 4.2 Global energy consumption by source (1965–2024)



Source: Adapted from Our World in Data, https://ourworldindata.org/grapher/energy-consumption-by-source-and-region, CC BY-SA 3.0

drawbacks of a renewable-heavy grid are becoming apparent.[5] They include the following:

- The cost of the dispatchable backup generation required to ensure high reliability (dispatchable refers to power sources that can be adjusted up or down on demand)
- The difficulties of synchronizing generators that lack mechanical inertia, which makes it harder to bring them online smoothly[6]
- The fire risks of grid-scale battery storage
- The critical materials required by clean energy technologies—materials that the United States heavily imports from countries whose interests do not always align with theirs (e.g., rare earths from China and cobalt from the Congo)

- **The difficulty of large-scale change in energy infrastructure** Energy is delivered to end users by systems, and those systems are hard to change, for fundamental reasons.[7] They involve large investments in assets that last decades, their parts need to work together (e.g., cars, fuel, and the fueling infrastructure must all be compatible), and there are many stakeholders whose interests often don't align. It also takes time to refine the hardware and operating procedures that ensure high reliability and efficiency. Energy systems are therefore best changed slowly and steadily over decades.

- **Efficiency limitations** Greater efficiency of end use (e.g., more miles per gallon in a vehicle or more lumens per watt in a light-emitting diode, or LED) is often invoked as an energy-saving measure. Yet such savings can be partially, or even totally, offset by direct rebound (i.e., greater efficiency leading to greater use) or indirect rebound (i.e., energy savings redirected to other uses).

- **Pragmatic challenges** Innovation in energy systems differs from other fields covered in the *Stanford Emerging Technology Review* (*SETR*) because many viable energy sources already exist. Any new energy source will be producing a commodity—fuel molecules or moving electrons (i.e., electricity). In this case, the cost of producing and delivering this commodity to the end user is paramount, subject to the other dimensions of the trilemma (reliability and lack of emissions). For large-scale deployment in a market economy, a new energy technology not only needs to work; it must be better than the alternatives. Moreover, in a globalized world, invention, manufacturing, and deployment often occur in different countries, which significantly affects the economic and security impacts of energy innovation.

In the United States, the Department of Energy (DOE) and the private sector have been the most significant funders of energy innovation, with government, academia, and the private sector conducting the research on energy technologies. Academia conducts the bulk of early-stage research on energy technologies, as do the US National Laboratories. However, these organizations don't have the resources to effect later-stage development or large-scale demonstrations, let alone deployments. Such efforts fall primarily to the private sector, which includes both large established companies and start-ups that might partner with academic institutions to take early-stage research to commercialization.

Key Developments

In the past few years, explosive growth in the demand for energy has begun to reshape the US energy economy. Data centers, artificial intelligence (AI) workloads, mining of cryptocurrencies, and industrial reshoring have driven electricity demand up by more than 2 percent annually for three consecutive years—a reversal of two decades of flat consumption. Forecasts suggest demand could rise between 16 and 25 percent over the next five years, straining grid capacity and prompting calls for new generation capacity.

In response, capital is shifting from intermittent renewables to dispatchable generation—especially natural gas, nuclear, and energy storage. Clean energy sectors are facing headwinds as utility-scale solar and wind projects have slowed due to the loss of federal tax credits and rising interest rates. Wind generation of electricity in the United States declined in 2023 for the first time in twenty-five years,[8] and growth in solar generation capacity slowed despite record installations in 2023.[9] At the same time, natural gas infrastructure is expanding rapidly, albeit constrained by labor and supply chain bottlenecks. Nuclear power is experiencing a renaissance, with new builds, restarts, and uprates underway. (Uprates refer to the process of increasing the maximum power output of an existing nuclear power plant through modifications or improvements.)

In short, aspirations for an accelerated energy transition have collided with scientific and techno-economic realities. These collisions, many of which predate the current US administration, have led to a new pragmatism in energy matters as the transition's costs and challenges become increasingly apparent. There is now more attention on an energy source being affordable and reliable than a single-minded focus on mitigating greenhouse gas emissions.[10]

This pragmatism is reflected in the repeal of most tax subsidies for emissions mitigation in the US Inflation Reduction Act and by the Trump administration's elevation of energy reliability and abundance, if not "energy dominance," over emissions-mitigation efforts. A number of other global developments during the past year reflect this trend as well:

- Global consumption of each of the major fossil fuels (coal, oil, and natural gas) hit a record high in 2024, despite record investments in renewable energy.[11] Total energy-related CO_2 emissions increased by 0.8 percent in 2024, hitting an all-time high of 37.8 gigatons of CO_2.[12]
- The United States began withdrawing from the Paris Agreement, while the European Union

Nuclear power is experiencing a renaissance, with new builds, restarts, and uprates underway.

softened compliance burdens under political pressure, signaling a retreat from an aggressive climate policy.[13]

- Mandates banning internal combustion engine vehicles are facing mounting resistance in Europe.[14] In addition, as Russia curtailed shipments of pipeline gas to Europe in retaliation for European support to Ukraine, Europe increased its imports of liquefied natural gas (LNG) to compensate for the loss.[15] However, LNG has a significantly worse emissions footprint than Russian pipeline gas because of the processing needed to liquify natural gas and the transportation of it from source to destination. Europe's actions therefore suggest a weakening of its commitments to reducing emissions, at least in the short term.

- Established in 2021 to align the financial sector with emissions targets, the Glasgow Financial Alliance for Net Zero unraveled in early 2025 following mass departures by major US banks from its affiliated coalitions, exposing the fragility of voluntary climate finance initiatives.[16] Companies are quietly retreating from public sustainability commitments amid a political backlash.[17] Major carmakers, including General Motors, Mercedes-Benz, and Aston Martin, scaled back EV plans in 2024–25, citing weak demand, high costs, and uncertain policy environments.[18]

The emerging zeitgeist is that all three legs of the trilemma are important—a change that will temper, but not halt, sustainable energy research and development and deployment efforts. In the past year, the two energy technologies that have gained in prominence are nuclear power and coal.

Nuclear Power

In September 2024, the DOE's "Pathways to Commercial Liftoff: Advanced Nuclear" report projected that the United States would need 700–900 gigawatts (GW) of clean firm power by 2050, with nuclear expected to triple its capacity to about 300 GW.[19] Consistent with this theme, the past year has seen a renaissance in the US nuclear power sector, driven by surging electricity demand, emissions concerns, and strategic industrial policy.

The most tangible development has been the completion of Vogtle Units 3 and 4 at the Alvin W. Vogtle Electric Generating Plant, in Georgia, which marked a historic milestone as the first new reactors built in the United States in over three decades (see figure 4.3).[20] These Westinghouse AP1000 pressurized water reactors added 2.2 GW of baseload capacity, making Plant Vogtle the largest nuclear power station in the country. Despite cost overruns and delays, the project demonstrated that large reactors remain viable when paired with federal loan guarantees and tax incentives.

Fermi America and Westinghouse have also announced plans to construct four AP1000 reactors near Amarillo, Texas, to power a massive AI data center.[21] The AP1000 design, now fully licensed and supported by a trained workforce and mature supply chain, is expected to see reduced costs and construction timelines compared to earlier builds. This proposal reflects a new trend of pairing nuclear power with energy-intensive digital infrastructure, as exemplified by Microsoft's proposed restart of a Three Mile Island unit to power its AI data centers.[22]

FIGURE 4.3 Vogtle Unit 3 under construction in October 2020

Source: US Nuclear Regulatory Commission

Holtec International received regulatory approval to restart the Palisades Nuclear Plant, in Michigan, which has been in decommissioning for three years.[23] Reactivations of a closed US nuclear plant like this could unlock latent capacity at other retired sites.

More technologically advanced reactors will generally use high-assay low-enriched uranium (HALEU), enriched to between 5 and 20 percent uranium-235. Yet domestic supply of this remains severely constrained. Historically, the United States has relied on Russian imports for HALEU, but recent legislation bans such imports after 2027, intensifying pressure to build a domestic supply chain.

Centrus Energy, the only US firm currently producing HALEU, delivered its first batch in late 2023 and aims to scale up its Ohio facility.[24] However, expansion will require billions of dollars in investment and sustained political support. In April 2025, the DOE awarded conditional HALEU supply commitments to five reactor developers by drawing from national stockpiles and DOE reserves.[25] To catalyze long-term supply, the DOE also awarded $2.7 billion in enrichment contracts to four firms, aiming to rebuild domestic enrichment capacity and reduce reliance on foreign sources.[26]

Coal

Coal-fired electricity in the United States has been declining since the 1950s, supplanted by inexpensive natural gas and the deployment of wind and solar generation.[27] Yet coal is garnering renewed attention as an option for powering data centers and AI that addresses reliability and supply chain vulnerability concerns. Coal is a domestically mined resource that can limit US dependence on foreign energy products.[28] In an effort to reinvigorate the industry, the US government designated coal as a critical mineral in April 2025, and President Trump signed an executive order, Reinvigorating America's Beautiful Clean Coal Industry, to expand its mining and use in the United States.[29]

Asia continues to be heavily dependent upon coal. China has pledged to be carbon neutral by 2060, and to that end, installed 365 GW of solar and wind energy capacity in 2024, far outpacing Europe.[30] Yet it also had almost 100 GW of new coal-fired capacity under construction in 2024, the greatest amount in a decade.[31] Coal also continues to dominate in India, accounting for 75 percent of generation.[32]

Over the Horizon

The discussion below provides an overview of technologies that are promising or where policy associated with them or an established area is less developed. It includes some energy technologies not covered in last year's edition of *SETR* and, due to space limitations, provides short descriptions of others. (For more on these technologies, refer to chapter 10 of *SETR* 2025.)

- **Thermal storage** Long-duration energy storage from intermittent sources such as wind and solar is necessary to capture their full value. Storage allows excess energy produced during plentiful times to be used when intermittent sources are unavailable. For example, solar energy is about twice as abundant in summer as in winter. Thermal storage stores excess power in the form of heat, such as by heating a large volume of salts to a very high temperature. When needed, this stored heat can be released to generate power.

- **Renewable combustible hydrocarbons and biodiesel** Research on these technologies aims to create energy sources that do not rely on fossil fuels such as oil, gas, or coal. Renewable fuels include combustible hydrocarbons such as biodiesel, which can be produced from animal fats or vegetable oils; bioethanol produced from corn or algae; hydrogen, which can be produced from many sources; and ammonia produced using green hydrogen (see definition of green hydrogen later in this chapter).

- **Carbon capture and storage (CCS)** CSS reduces CO_2 in the atmosphere by capturing and storing it underground. Carbon capture is usually done at the emission source, like power plant smokestacks, using materials and membranes to extract CO_2 for underground storage or material use. Direct air capture, which removes CO_2 from the atmosphere at much lower concentrations, consumes more energy. However, it is currently progressing through several commercial-scale projects, with research focused on scaling, storage duration, and cost-effectiveness of various removal methods such as biomass storage and mineralization.

- **New grid technologies** These are needed to manage the future electric grid. Compared to today's grid, this future one will be larger, more complex, and more decentralized, integrating varied renewable energy sources and dealing with increased electricity demand. Relevant new technologies include reconductoring (replacing existing power cables with more advanced ones) to boost line capacity, end-use energy management to shift and optimize the timing of electricity consumption, vehicle-to-grid systems that allow EVs to feed energy back to the grid, and second-life battery applications for stationary storage.

Additionally, AI and data-driven systems will optimize grid operations and maintenance by responding dynamically to changes in renewable generation and demand, and by predicting equipment failures.

Of particular interest this year are the following technologies, for which longer descriptions are offered:

Green Hydrogen

Hydrogen is vital to today's energy systems, with important roles in refining and in fertilizer and steel production. Many envision it to be similarly vital in a deeply decarbonized energy future, either as a vehicle fuel itself or as a component of a synthetic fuel, as a form of grid-scale storage, and as an input to industrial processes.

Hydrogen today is produced almost exclusively by steam reforming, in which methane and water are passed over catalysts at suitable temperatures and pressures to yield hydrogen and CO_2.[33] In a decarbonized world, that process would be replaced by the electrolysis of water driven by carbon-free electricity (e.g., from renewables). While there have been initial steps in that direction, progress will depend upon the cost of "green" hydrogen (i.e., hydrogen produced through renewable electricity sources) relative to that produced by steam reforming, the scalability of the technology, and the extent to which the world pursues decarbonization.

Electrolyzer technologies—primarily alkaline, proton exchange membrane, and emerging solid oxide systems—have seen dramatic cost reductions over the past decade. DOE estimates suggest electrolyzer costs have dropped by over 80 percent since 2005, with further reductions expected as manufacturing scales and efficiency improves.[34] Global installed electrolyzer capacity now exceeds 4.5 GW, up from just 0.17 GW in 2021. But despite these advances and a 10 percent annual growth in demand, green hydrogen still accounted for less than 1 percent of the roughly 100 million tons of global hydrogen production in 2024.

While political changes have reduced US prospects and enthusiasm for green hydrogen, global investment is growing. The International Energy Agency projected $7.8 billion in global spending on clean hydrogen in 2025—a 70 percent increase over 2024—with 6 GW allocated to electrolysis projects. Countries like India, China, and Oman are launching giga-scale projects and positioning themselves as future exporters.

The extent to which those investments will pay off depends upon several factors:

- **Cost** Green hydrogen is still expensive and capital-intensive. Today's cost, $3 to $6 per kilogram, remains significantly higher than the $1 to $2 per kilogram cost of steam reforming. Achieving cost parity by 2031 will require breakthroughs in efficiency, materials, and manufacturing.
- **Infrastructure** Pipelines, storage, and refueling stations are limited, especially outside pilot hubs.
- **Offtake uncertainty** Many projects lack firm buyers, creating a mismatch between final investment decisions and market demand.
- **Regulatory fragmentation** Inconsistent standards for emissions accounting, certification, and trade hinder the formation of a global hydrogen market.

Geothermal Energy

Geothermal energy draws upon Earth's internal heat, making it, unlike wind and solar technologies, independent of weather conditions. The most common method of extracting geothermal energy

Compared to today's grid, [the] future one will be larger, more complex, and more decentralized.

FIGURE 4.4 Drone image of the setup for the first demonstration of Quaise Energy's novel drilling technique on a full-scale oil rig

Source: Quaise Energy

is to use steam from naturally occurring geysers to generate electricity. The hot water from geothermal springs can also be used to heat buildings directly.[35] However, there are a limited number of sites where these methods can be deployed.

An alternative method for extracting geothermal energy is to inject water into dry, hot rock deeper down and bring the resulting steam to the surface to generate electricity.[36] Beyond knowing where best to drill, the use of standard drills limits the depths that can be reached and thus the temperatures that can be accessed. And because fracking (which involves injecting liquid at high pressure to expand existing fissures) is required to create channels for steam generation, it is possible to induce seismic events that can damage local infrastructure.[37]

To address these issues, various start-ups are developing safe and efficient drilling techniques that can access greater depths. Quaise Energy, for example, has recently developed a new drill that uses an electromagnetic beam to vaporize rock at great depths (see figure 4.4),[38] thus addressing many of the challenges faced by physical drills.[39] The deepest hole ever drilled was almost eight miles deep, but Quaise wants to push to twelve miles.

There are also several recent innovations in dry heat geothermal generation and storage, which increase the flexibility and availability of geothermal energy. For example, Fervo, a start-up based out of Houston, drills horizontally to find heat sources that can warm the water that it injects.[40] Oil and gas companies have become increasingly interested in this form of technology because oil wells can be retrofitted for geothermal energy.[41]

One generic challenge in geothermal applications is depletion of the heat resource because of the long amount of time it takes for the ambient underground heat to replace that which has been extracted. For example, the capacity of the Geysers geothermal plants in California has declined some 65 percent in the past three decades, from 2,000 megawatts (MW) in 1987 to 725 MW today.[42]

The National Renewable Energy Laboratory estimates that geothermal energy on federal lands could support as much as 975 GW of dispatchable generation.[43] The technology has significant bipartisan support. The Geothermal Energy Opportunity Act was introduced in Congress in January 2025 to accelerate approval of geothermal projects.[44] On May 30, 2025, the US Department of the Interior announced emergency permitting procedures to accelerate geothermal projects in support of the Trump administration's goal of energy dominance.[45]

Small Modular Nuclear Reactors

A small modular reactor (SMR) is a compact nuclear fission reactor with an electric power output of up to 300 MW. It is designed for factory fabrication and modular installation to provide a flexible, scalable low-carbon energy source. More than a dozen US start-ups are racing to commercialize advanced SMR designs that promise lower costs, faster deployment, and enhanced safety. These include molten salt, liquid metal, and high-temperature, gas-cooled reactors, many of which operate at atmospheric pressure and use passive safety systems. (Passive safety systems in nuclear reactors enhance safety by relying on natural physical phenomena—such as gravity, natural circulation, and convection—to maintain safe reactor conditions without external power, active controls, or operator intervention.)

Among them are:

- Kairos Power, which is developing the Hermes 2 reactor in Tennessee.[46] This uses molten fluoride salt as coolant and is expected to deliver 50 MW to Google's data centers by 2030, with plans to scale to 500 MW by 2035.
- TerraPower, which broke ground on its Natrium reactor in Wyoming.[47] Using liquid sodium coolant and molten salt energy storage, the plant aims to dispatch 345 MW by 2030 to PacifiCorp, with ramp-up capability to 500 MW.
- Oklo, which is pursuing compact fast reactors cooled by liquid metal.[48] Despite regulatory setbacks, it has secured agreements to supply 12 GW to data center operator Switch by 2044.

Reflecting the urgency to meet AI-driven electricity demand, the DOE recently selected eleven projects for a pilot program to fast-track the development of advanced reactor technologies. Some of these will be intended for use in small modular test reactors, with the aim to bring at least three online by July 4, 2026.[49] One of these, Deep Fission, plans to put an SMR in a one-mile deep borehole that will provide much of the physical safety barriers needed for reactors and thereby reduce construction costs significantly.[50]

Fusion Energy

Like the energy produced by splitting heavy atoms such as uranium or plutonium (fission), energy produced by combining two light atoms (fusion) entails no greenhouse gas emissions. It also has the additional advantages of using abundant fuels and producing minimal long-lived radioactive waste.[51] While it is still some distance from commercial demonstration, nuclear fusion could prove to be a viable long-term energy source for future generations.

The most important among the several technical challenges to realizing fusion energy is to achieve

Many nuclear fusion advances are now driven by start-ups, reflecting the private sector's central role in energy innovation.

"gain"—that is, to confine a plasma of hydrogen isotopes for durations and at temperatures and densities sufficient to produce more energy from fusion than was required to create the plasma. One approach to solving this confinement problem is magnetic confinement fusion, which uses powerful magnets to contain and control a superheated plasma of deuterium and tritium. A second is inertial confinement fusion, which calls for rapidly compressing a deuterium-tritium fuel pellet using lasers to ignite the fusion reaction.

In 2022, the National Ignition Facility, which has the world's largest laser, achieved the first laboratory fusion system with net gain—that is, it produced fusion energy 1.5 times greater than the laser energy that created the hot, dense mass of hydrogen. Subsequent refinement led to an April 2025 experiment that showed a gain of 4.2 times.[52] Although this is a promising milestone, a gain of mid-double-digits magnitude is necessary for a viable power plant.

Beyond sustaining a plasma of sufficient gain, there are two other major technological challenges in making fusion a viable source of energy.

- The walls of whatever vessel contains the plasma must be robust. A fusion plasma produces X-rays and particles that will rapidly degrade wall material. It's therefore important to find material that can resist (or at least slow) this degradation.
- The first fusion reactors will almost certainly be fueled by a mixture of deuterium and tritium. While deuterium is readily available in nature, tritium is radioactive, with a half-life of 12.3 years. This means that tritium must be manufactured. A fusion power plant with a 1 GW output operating for a year would consume at least several times the current global production of tritium. Therefore, a viable fusion reactor must breed its own tritium. Self-manufacture of tritium can be done in principle by exposing ^{6}Li (a particular and relatively rare isotope of lithium) to the neutrons that the reactor produces. However, this process has never been demonstrated at the scale required.

Even after these hurdles are surmounted, the cost of electricity generated must be competitive with conventional alternatives. Most knowledgeable observers believe fusion power into the grid won't happen until 2040, at the earliest.

France, Japan, and China are all making progress in their national programs pursuing magnetic confinement fusion.[53] America's national program is dominated by participation in the ITER tokamak international initiative to build a fusion reactor, although the DOE has recently reinvigorated efforts on the alternative stellarator concept. (A tokamak is easier to build and more efficient than a stellarator design, but a tokamak must operate in pulses rather than continuously. A stellarator can operate continuously but is more difficult to build.)

Many nuclear fusion advances are now driven by start-ups, reflecting the private sector's central role in energy innovation.[54] Private companies are pursuing a wider range of fusion technologies than government-backed programs, across both inertial and magnetic confinement approaches. In magnetic confinement, notable examples include Commonwealth Fusion Systems (originating from MIT research)[55] and TAE Technologies (founded at University of California–Irvine),[56] which plan to build their first fusion reactors by the early 2030s. German start-up Proxima Fusion has released open-source plans for a nuclear fusion power plant employing novel containment strategies.[57]

Iron-Air Batteries

Storing energy in large-scale, long-duration batteries is one way of compensating for intermittent wind and solar generation. Unfortunately, lithium-ion (Li-ion) batteries are ill suited for this purpose due to their limited storage capacity and high cost. In *SETR* 2025, we highlighted novel batteries that are emerging to meet future energy reliability needs, including redox flow, Ni-H_2 gas, and Zn-MnO_2. In the past year, Form Energy's iron-air batteries have also emerged as a promising alternative.[58]

Iron-air batteries can deliver up to one hundred hours of utility-scale storage, compared to four hours from Li-ion batteries. (The time of utility-scale storage refers to the period over which a battery can sustain its full rated power output.) This improvement stems from a reversible rusting process. To discharge an iron-air battery, iron oxidizes (i.e., it rusts by combining with oxygen), causing a flow of electrons (electricity). To charge the battery, electricity is used to reverse the rusting process, releasing oxygen.[59] While the chemical processes in Li-ion batteries allow for rapid charge and discharge cycles, they also lead to faster battery degradation. In contrast, the reversible rusting process enables extended energy release, making iron-air batteries ideal for long-duration storage.

Because iron is one of the most abundant metals on Earth, the batteries are significantly cheaper than Li-ion or redox flow batteries. They are also safer, with no flammable materials.[60] But they are less efficient than traditional batteries, releasing only 50 to 60 percent of the stored energy compared to 80 to 90 percent for Li-ion batteries.[61] Even so, many utility companies plan to use this novel energy storage technology. PacifiCorp, for example, noted in its 2025 integrated resource plan that "100-hour iron-air storage has a low capital cost with a low round trip efficiency," making it "a valuable asset in its portfolio."[62]

Iron-air batteries are not intended to completely replace Li-ion batteries but rather serve as a complementary form of long-duration energy storage. They will also contribute to national security goals by reducing US dependence on Chinese-dominated Li-ion batteries.[63]

Transportation Electrification

The One Big Beautiful Bill that passed in 2025 eliminated the tax incentives for EVs and the associated charging infrastructure that had been in place under the 2022 Inflation Reduction Act. This will slow EV adoption in the United States. Nevertheless, the electrification of light-duty transport continues apace in other parts of the world.[64] (Light-duty transport refers to vehicles designed primarily for the transportation of passengers or cargo and weighing less than 8,500 pounds. It typically includes passenger cars, small vans, SUVs, and pickup trucks.)

Certain EV technologies, such as batteries, are currently dominated by Chinese suppliers, who lead the global market. For example, Contemporary Amperex Technology Co., Limited (CATL), supplies 35 percent of the world's Li-ion EV batteries, serving major companies including Tesla, BMW, and Volkswagen. In early May 2025, CATL raised $4.6 billion in a Hong Kong IPO, notably excluding American investors.[65]

Range is one of the greatest challenges facing EVs, making improving battery technology key to their adoption. Currently, EVs with fast charging speeds and long ranges, such as the Tesla Model Y and Mercedes-Benz EQS, take roughly 30 minutes to charge to a range of 280 miles. However, in April 2025, CATL announced that its latest EV battery could add 323 miles of driving range with 5 minutes of charging, marking a significant improvement in EV battery range and charging speed.[66] However, the implications for charging infrastructure should not be underestimated—assuming a nominal 3 miles per kWh, adding 300 miles of range in 5 minutes requires 1.2 MW of power at the charging station, or about 1,000 times the average power draw of an American household.

Policy Issues

The past year has seen dramatic changes in US energy policy, regulation, and economics, driven by surging electricity demand, shifting political priorities, and global market pressures. Federal and state governments have enacted sweeping reforms, regulatory agencies are reorienting their frameworks, and economic forces are reshaping investment flows. The following are some of the key policy-related trends:

From Decarbonization to Energy Dominance

The Trump administration has emphasized energy abundance and industrial competitiveness over greenhouse gas mitigation, reversing many of the clean energy priorities of previous years. This shift is embodied in legislation that eliminates tax credits for new wind and solar projects—including the Production Tax Credit and Investment Tax Credit, which had catalyzed renewable deployment for decades. New legislation has also expanded federal support for nuclear power, geothermal, and natural gas infrastructure, including fast-track permitting and loan guarantees.

At the same time, the administration issued four executive orders aimed at "unleashing American energy," including mandates to accelerate fossil fuel leasing on federal lands, streamline pipeline approvals, and sunset outdated regulations.[67] These moves signal a realignment toward energy security, grid reliability, and domestic production.

Regulatory Overhaul

Regulatory agencies have undergone significant restructuring. The DOE launched the largest deregulatory effort in its history, proposing to eliminate or modify forty-seven regulations, ranging from appliance-efficiency standards to environmental review procedures. According to the DOE,[68] these changes are expected to save consumers an estimated $11 billion and reduce regulatory text by over 125,000 words.

The Federal Energy Regulatory Commission was directed to implement "conditional sunset clauses" for all energy-related regulations, requiring periodic review and expiration of such regulations unless reauthorized. This introduces uncertainty into long-standing rules governing transmission planning, interconnection (i.e., links between local grids), and wholesale market operations.

The Environmental Protection Agency (EPA) scaled back enforcement of greenhouse gas emissions standards for power plants and vehicles. While previous rules aimed to reduce emissions through 2032, new guidance allows states greater flexibility and delays compliance timelines. More importantly, the EPA has also proposed rescinding the Endangerment Finding that underpins its regulation of greenhouse gas emissions.[69] As noted earlier, nuclear power is enjoying a revival that spans large-scale reactor projects, a host of SMR start-ups, renewed attention to advanced designs and advanced fuels, and federal initiatives aimed at rebuilding domestic capacity and global leadership. For example, the Nuclear Regulatory Commission (NRC) has accelerated licensing pathways for SMRs and microreactors (i.e., advanced reactors generating no more than about a tenth of the power of an SMR and designed for mobility, fast deployment, and minimal onsite staffing), reflecting bipartisan support for advanced nuclear technologies. Additionally, in May 2025 the White House issued four executive orders to reinvigorate the nuclear industrial base, streamline reactor licensing, reform the NRC, and deploy advanced reactors for national security.[70] These orders mandate:

- Accelerated licensing for SMRs and microreactors
- Funding to restart closed plants and uprate existing reactors
- Support for ten new large reactor designs under construction by 2030
- Expansion of HALEU enrichment and deconversion infrastructure
- Workforce development and supply chain localization

Congress has also passed the Accelerating Deployment of Versatile, Advanced Nuclear for Clean Energy (ADVANCE) Act, which facilitates reactor deployment and strengthens export capabilities.[71] The latter is needed to compete with aggressive Chinese and Russian expansion of their nuclear exports.

Supply Chain Issues

Access to lithium and other critical minerals poses a challenge to the long-term success of domestic Li-ion EV battery production. China currently produces 90 percent of the world's permanent rare-earth magnets, which are critical components in EV motors. Since December 2025, China has imposed regulations on exporting various critical minerals, which could gravely impact American access to the materials necessary to build EVs and many other sustainable technologies.[72] Many companies are also concerned about the long-term availability of these materials, and some, like CATL, are looking into more efficient ways to recycle old Li-ion batteries.[73] The battery recycling market is predicted to grow at a compound annual growth rate of 40 percent.

NOTES

1. "International Energy Outlook—Consumption," US Energy Information Administration, October 2021, https://www.eia.gov/outlooks/ieo/consumption/sub-topic-03.php.

2. Mark Schipper and Tyler Hodge, "After More Than a Decade of Little Change, U.S. Electricity Consumption Is Rising Again," *US Energy Information Administration*, May 13, 2025, https://www.eia.gov/todayinenergy/detail.php?id=65264.

3. Energy Institute, "Statistical Review of World Energy 2025," with major processing by Our World in Data, "Other Renewables (Including Geothermal and Biomass)" [dataset], https://ourworldindata.org/grapher/energy-consumption-by-source-and-country?overlay=sources.

4. Euan Graham and Nicolas Fulghum, *Global Electricity Review 2025*, Ember Energy, April 8, 2025, https://ember-energy.org/latest-insights/global-electricity-review-2025.

5. "Department of Energy Releases Report on Evaluating U.S. Grid Reliability and Security," US Department of Energy, July 7, 2025, https://www.energy.gov/articles/department-energy-releases-report-evaluating-us-grid-reliability-and-security.

6. Mechanical inertia in an electrical generation system is vital because it resists sudden changes in the rotational speed of generators, thereby smoothing frequency fluctuations resulting from variable loads and providing critical time for other control systems to stabilize the power grid and maintain consistent electricity supply. A lack of sufficient mechanical interia contributed to a major grid failure on the Iberian Peninsula in April 2025, according to a report by the Baker Institute (see Raúl Bajo Buenestado, "The Iberian Peninsula Blackout—Causes, Consequences, and Challenges Ahead," Baker Institute, May 2, 2025, https://www.bakerinstitute.org/research/iberian-peninsula-blackout-causes-consequences-and-challenges-ahead).

7. Steven Koonin and Avi Gopstein, "Accelerating the Pace of Energy Change," *Issues in Science and Technology*, 2011, https://issues.org/koonin/.

8. "Wind Generation Declined in 2023 for the First Time Since the 1990s," US Energy Information Administration, April 30, 2024, https://www.eia.gov/todayinenergy/detail.php?id=61943.

9. "Global Market Outlook for Solar Power 2025–2029," *SolarPower Europe*, 2025, https://www.solarpowereurope.org/insights/outlooks/global-market-outlook-for-solar-power-2025-2029/detail.

10. Daniel Yergin, Peter Orszag, and Atul Arya, "The Troubled Energy Transition," *Foreign Affairs*, February 25, 2025, https://www.foreignaffairs.com/united-states/troubled-energy-transition-yergin-orszag-arya.

11. "Home," "2025 Statistical Review of World Energy," Energy Institute, https://www.energyinst.org/statistical-review/home.

12. "CO2 Emissions—Global Energy Review 2025," International Energy Agency, 2025, https://www.iea.org/reports/global-energy-review-2025/co2-emissions.

13. "Sustainability Policy Split Widens Between U.S. and EU in 2025," KnowESG, August 7, 2025, https://www.knowesg.com/regulators/sustainability-policy-split-widens-between-u-s-and-eu-in-2025.

14. Jordyn Dahl, Giorgio Leali, and Oliver Noyan, "The EU Ban on Combustion Car Engines Is in Trouble," *Politico*, March 11, 2025,

https://www.politico.eu/article/why-eu-combustion-car-ban-is-in-trouble-greenhouse-gas-climate-change/.

15. "Security of Gas Supply," European Commission, 2022, https://energy.ec.europa.eu/topics/energy-security/security-gas-supply_en.

16. Gurjinder Khambay, "GFANZ Is in Freefall: So What Happens Next?," *Finance Innovation Lab*, January 20, 2025, https://financeinnovationlab.org/gfanz-is-in-freefall-so-what-happens-next/.

17. ENH Curators, "Corporate Climate Promises Are Collapsing as Companies Retreat from Green Goals," *Environmental Health News*, June 13, 2025, https://www.ehn.org/corporate-climate-promises-are-collapsing-as-companies-retreat-from-green-goals.

18. Anthony Capretto, "All the Automakers That Have Backpedaled or Pushed Back EV Production," *CarBuzz*, August 3, 2024, https://carbuzz.com/all-the-automakers-that-have-pushed-back-ev-production/.

19. *Pathways to Commercial Liftoff: Advanced Nuclear*, US Department of Energy (DOE), September 2024. This report is no longer available on the DOE website but can be found elsewhere (e.g., https://cdn.prod.website-files.com/650ac01414c2440d4b402689/67ea82f863d6e92b674683aa_U.S.%20Department%20of%20Energy%20-%20Pathways%20to%20%20Commercial%20Liftoff%20-%20Advanced%20Nuclear.pdf).

20. Slade Johnson, "Plant Vogtle Unit 4 Begins Commercial Operation," *US Energy Information Administration*, May 1, 2024, https://www.eia.gov/todayinenergy/detail.php?id=61963.

21. "Fermi America Partners with Westinghouse to Support Licensing for Four AP1000 Units," Westinghouse Electric Company, August 21, 2025, https://info.westinghousenuclear.com/news/fermi-america-partners-with-westinghouse-to-support-licensing-for-four-ap1000-units.

22. Jericho Casper, "Three Mile Island Nuclear Plant May Return to Fuel Microsoft AI Data Center," *Broadband Breakfast*, March 27, 2025, https://broadbandbreakfast.com/three-mile-island-nuclear-plant-may-return-to-fuel-microsoft-ai-data-center/.

23. Amy Grant, "NRC Reauthorizes Palisades Operating License in Historic First for U.S. Nuclear Industry," *Holtec International*, July 24, 2025, https://holtecinternational.com/2025/07/24/hh-40-15/.

24. "Home," Centrus Energy Corp., June 20, 2025, https://www.centrusenergy.com/.

25. "U.S. Department of Energy to Distribute First Amounts of HALEU to U.S. Advanced Reactor Developers," US Department of Energy, April 9, 2025, https://www.energy.gov/articles/us-department-energy-distribute-first-amounts-haleu-us-advanced-reactor-developers.

26. "DOE Names Four Companies to Split $2.7 Billion in Future HALEU Enrichment Contracts," Nuclear Newswire, 2025, https://www.ans.org/news/article-6485/doe-names-four-haleu-enrichment-contractors/.

27. Charles Kolstad, *What Is Killing the U.S. Coal Industry*?, Stanford Institute for Economic Policy Research, Stanford University, March 2017, https://siepr.stanford.edu/publications/policy-brief/what-killing-us-coal-industry.

28. Antonio Olivo, "Internet Data Centers Are Fueling Drive to Old Power Source: Coal," *Washington Post*, April 17, 2024, https://www.washingtonpost.com/business/interactive/2024/data-centers-internet-power-source-coal/.

29. "Reinvigorating America's Beautiful Clean Coal Industry and Amending Executive Order 14241," The White House, April 8, 2025, https://www.whitehouse.gov/presidential-actions/2025/04/reinvigorating-americas-beautiful-clean-coal-industry-and-amending-executive-order-14241/.

30. Qi Qin and Christine Shearer, "When Coal Won't Step Aside: The Challenge of Scaling Clean Energy in China," *Centre for Research on Energy and Clean Air*, April 21, 2025, https://energyandcleanair.org/publication/when-coal-wont-step-aside-the-challenge-of-scaling-clean-energy-in-china/.

31. Anika Patel, "China's Construction of New Coal-Power Plants Reached 10-Year High in 2024," *Carbon Brief*, February 13, 2025, https://www.carbonbrief.org/chinas-construction-of-new-coal-power-plants-reached-10-year-high-in-2024/.

32. Sibi Arasu, "The World's Most Populous Nation Is Making Big Strides in Green Energy Transition," *Associated Press News*, June 2025, https://apnews.com/article/climate-change-india-renewable-solar-coal-wind-power-ffaaa2446482f0b96516045528ed690b.

33. Farid Safari and Ibrahim Dincer, "A Review and Comparative Evaluation of Thermochemical Water Splitting Cycles for Hydrogen Production," *Energy Conversion and Management* 205 (December 2019): 112182, https://doi.org/10.1016/j.enconman.2019.112182.

34. "Progress in Hydrogen and Fuel Cells," US Department of Energy Office of Energy Efficiency and Renewable Energy, January 2025, https://www.energy.gov/sites/default/files/2025-01/progress-hydrogen-fuel-cells-jan2025.pdf.

35. "Geothermal Basics," US Department of Energy, 2019, https://www.energy.gov/eere/geothermal/geothermal-basics.

36. "Enhanced Geothermal Systems," US Department of Energy, 2023, https://www.energygov/eere/geothermal/enhanced-geothermal-systems.

37. Jules Bernstein, "Study Ties Fracking to Another Type of Shaking," *University of California, Riverside News*, August 10, 2023, https://news.ucr.edu/articles/2023/08/10/study-ties-fracking-another-type-shaking.

38. Benoît Morenne, "Can a Geothermal Startup Vaporize Rock to Drill the Deepest Holes Ever?," *Wall Street Journal*, March 6, 2025, https://www.wsj.com/business/energy-oil/can-a-geothermal-startup-vaporize-rock-to-drill-the-deepest-holes-ever-9f1e3c2d.

39. Joe Salas, "Quaise Demos Maser Drill Bit to Go Deeper Than Humans Have Ever Gone," *New Atlas*, May 29, 2025, https://newatlas.com/energy/quaise-energy-millimeter-wave-drill-demo-houston/.

40. Brad Plumer, "Hungry for Clean Energy, Facebook Looks to a New Type of Geothermal," *New York Times*, August 26, 2024, https://www.nytimes.com/2024/08/26/climate/meta-facebook-geothermal-fracking-energy.html.

41. Alejandro De La Garza and Andrew D. Johnson, "Geothermal Energy Could Be Huge, But We Need Oil and Gas Companies to Build It," *Time*, November 22, 2023, https://time.com/6338438/oil-gas-geothermal-energy/.

42. Saul Elbein, "Pilot Project Seeks to Fix Achilles' Heel of Geothermal Power," *The Hill*, May 8, 2025, https://thehill.com/policy/energy-environment/5288324-pilot-project-geothermal-energy-california/.

43. "New Interagency Study Finds Further Expansion of Renewable Energy Production on Federal Lands Could Power Millions

More American Homes by 2035," US Department of Energy, January 14, 2025, https://www.energy.gov/articles/new-interagency-study-finds-further-expansion-renewable-energy-production-federal-lands.

44. Clara Hudson, "Republicans Push for Surge in Geothermal Energy on Federal Land," *Wall Street Journal*, May 13, 2025, https://www.wsj.com/articles/republicans-push-for-surge-in-geothermal-energy-on-federal-land-79f27914.

45. "Department of the Interior Implements Emergency Permitting Procedures to Accelerate Geothermal Energy Development for National Security and Energy Independence," US Department of the Interior, May 30, 2025, https://www.doi.gov/pressreleases/department-interior-implements-emergency-permitting-procedures-accelerate-geothermal.

46. "Homepage," Kairos Power, 2025, https://kairospower.com/.

47. "Natrium Nuclear Energy | Isotopes | Cancer Treatment," TerraPower, 2025, https://www.terrapower.com/.

48. "Home," Oklo Inc., 2025, https://www.oklo.com/overview/default.aspx.

49. "Department of Energy Announces Initial Selections for New Reactor Pilot Program," US Department of Energy, August 12, 2025, https://www.energy.gov/articles/department-energy-announces-initial-selections-new-reactor-pilot-program.

50. "Deep Fission Nuclear Energy Solutions," Deep Fission Inc., August 2025, https://deepfission.com/.

51. "Advantages of Fusion," ITER Organization, 2023, https://www.iter.org/fusion-energy/advantages-fusion.

52. Skye Jacobs, "Fusion Breakthrough: NIF Achieves 8.6 Megajoules, Shattering Previous Record," *TechSpot*, May 19, 2025, https://www.techspot.com/news/107971-fusion-breakthrough-nif-achieves-86-megajoules-shattering-previous.html.

53. On France: Bob Rubila, "France Could Become the First Nation to Master Nuclear Fusion, the Ultimate Dream of Humanity," *Farmingdale Observer*, May 20, 2025, https://farmingdale-observer.com/2025/05/20/france-could-become-the-first-nation-to-master-nuclear-fusion-the-ultimate-dream-of-humanity/; on Japan: Ryuto Imao and Kento Fukui, "Japan Plans to Launch Pilot Fusion Power Plant Next Decade," *Nikkei Asia*, May 20, 2025, https://asia.nikkei.com/business/energy/japan-plans-to-launch-pilot-fusion-power-plant-next-decade; on China: Bob Mumgaard, "China Just Bet $2 Billion on Fusion Energy. The U.S. Must Respond," *The Hill*, August 2025, https://thehill.com/opinion/energy-environment/5431531-fusion-energy-china-us-race/.

54. Raymond Zhong, "Inching Toward a Fusion Energy Future," *New York Times*, November 19, 2024, https://www.nytimes.com/2024/11/19/climate/fusion-energy-startups.html.

55. "Home," Commonwealth Fusion Systems, 2024, https://cfs.energy/.

56. "Clean Energy Solutions for a Bright Future," TAE Technologies, April 15, 2025, https://tae.com/.

57. Yusuf Khan, "German Startup Publishes Open-Source Plans for Nuclear-Fusion Power Plant," *Wall Street Journal*, February 26, 2025, https://www.wsj.com/articles/german-startup-publishes-open-source-plans-for-nuclear-fusion-power-plant-7b2b6241.

58. Scott J. Mulligan, "2024 Climate Tech Companies to Watch: Form Energy and Its Iron Batteries," *MIT Technology Review*, October 2024, https://www.technologyreview.com/2024/10/01/1104382/2024-climate-tech-companies-form-energy-iron-batteries/.

59. Matt Blois, "The Search for Long-Duration Energy Storage," *C&EN Global Enterprise* 103, no. 5 (2025): 30–35, https://doi.org/10.1021/cen-10305-cover.

60. Jan Ing Girschik, "ELuStat: Iron-Air Battery as Stationary Energy Storage," *Fraunhofer Institute for Environmental, Safety and Energy Technology UMSICHT*, 2019, https://www.umsicht.fraunhofer.de/en/projects/iron-air-battery.html.

61. Staff writer, "Will Iron-Air Batteries Revolutionize Renewable Energy Storage?," *Environment + Energy Leader*, August 19, 2024, https://www.environmentenergyleader.com/stories/will-iron-air-batteries-revolutionize-renewable-energy-storage,48339.

62. April Bonner, "PacifiCorp Looks to Add 3,073 MW of Multi-Day Duration Iron-Air Battery Storage in 2025 IRP," *Energy-Storage.News*, April 10, 2025, https://www.energy-storage.news/pacificorp-looks-to-add-3073mw-of-iron-air-storage-in-2025-irp/.

63. Ellen Wald, "The U.S. Wants to End Its Reliance on Chinese Lithium. Its Policies Are Doing the Opposite," *Atlantic Council*, January 23, 2024, https://www.atlanticcouncil.org/blogs/new-atlanticist/the-us-wants-to-end-its-reliance-on-chinese-lithium-its-policies-are-doing-the-opposite/.

64. "Trends in Electric Car Markets—Global EV Outlook 2025," International Energy Agency, 2025, https://www.iea.org/reports/global-ev-outlook-2025/trends-in-electric-car-markets-2.

65. Jennifer Jett and Eve Qiao, "China's EV Battery Leader Surges in World's Biggest Listing This Year," *NBC News*, May 20, 2025, https://www.nbcnews.com/world/asia/china-ev-battery-electic-vehicle-listing-catl-rcna207619.

66. Dylan Butts, "China's CATL Claims to Beat BYD's EV Battery Record with Longer Range on a 5-Minute Charge," *CNBC*, April 22, 2025, https://www.cnbc.com/2025/04/22/chinas-catl-claims-to-beat-byds-ev-battery-record-with-longer-range-on-a-5-minute-charge.html.

67. "Unleashing American Energy," The White House, January 21, 2025, https://www.whitehouse.gov/presidential-actions/2025/01/unleashing-american-energy/.

68. "Energy Department Slashes 47 Burdensome and Costly Regulations, Delivering First Milestone in America's Biggest Deregulatory Effort," US Department of Energy, May 13, 2025, https://www.energy.gov/articles/energy-department-slashes-47-burdensome-and-costly-regulations-delivering-first-milestone.

69. "Proposed Rule: Reconsideration of 2009 Endangerment Finding and Greenhouse Gas Vehicle Standards," US Environmental Protection Agency, May 13, 2025, https://www.epa.gov/regulations-emissions-vehicles-and-engines/proposed-rule-reconsideration-2009-endangerment-finding.

70. Andy Kriha, Jason Hill, Kenneth Cestari, Taite McDonald, and Elizabeth Noll, "President Trump Signs 4 Executive Orders to Deploy New Nuclear Reactors, Strengthen Supply Chain," *Holland & Knight Insights*, May 28, 2025, https://www.hklaw.com/en/insights/publications/2025/05/president-trump-signs-4-executive-orders.

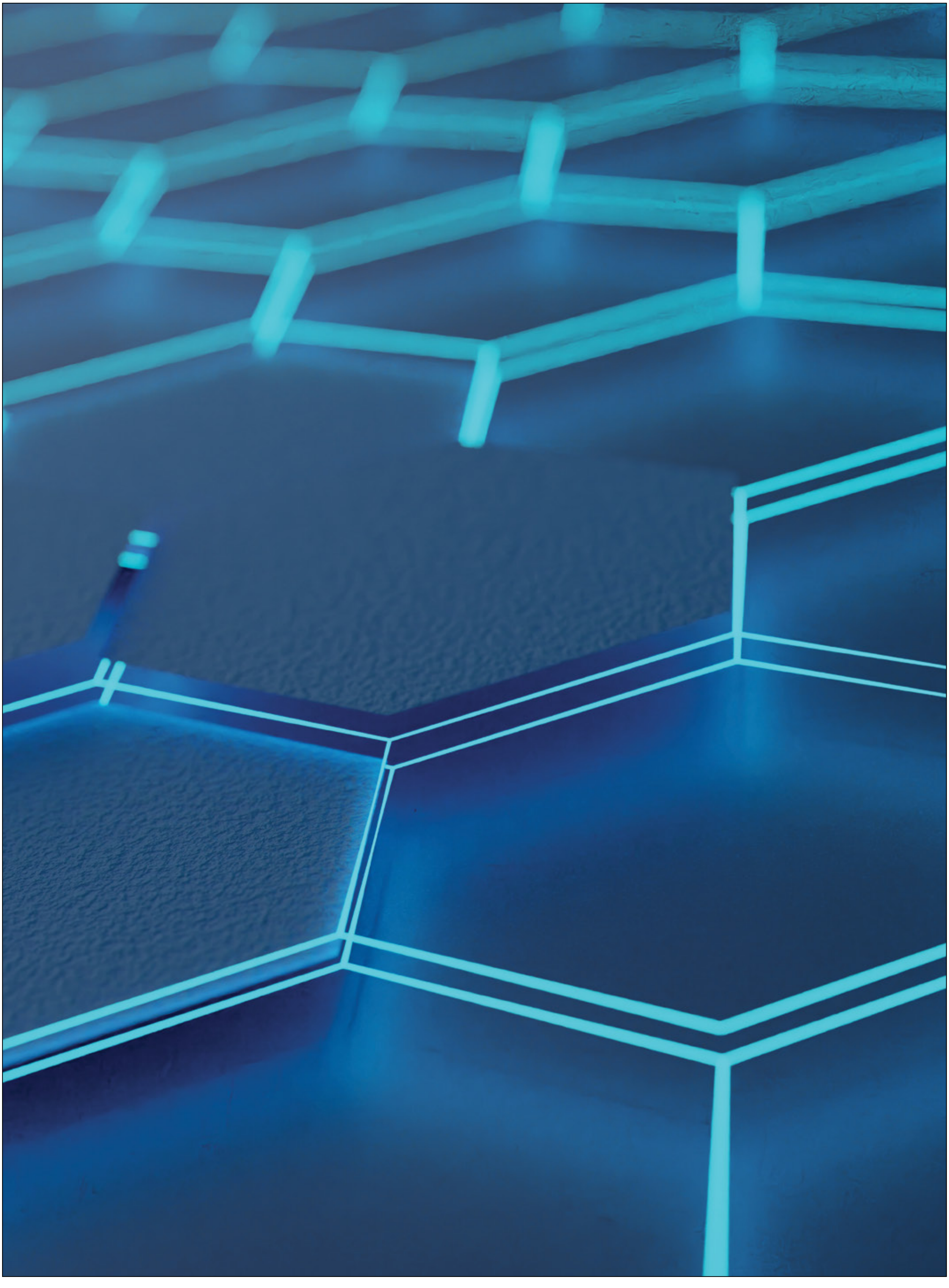
71. Justine Calma, "Congress Votes to Advance Nuclear Energy Development in the U.S.," *The Verge*, June 19, 2024, https://www.theverge.com/2024/6/19/24181808/congress-pass-nuclear-energy-bill-biden-signature.

72. Mia Nulimaimaiti, "China's Critical Metal Exports Plummet as Trade War Curbs Take Effect," *South China Morning Post*, April 22, 2025, https://www.scmp.com/economy/china-economy/article/3307410/chinas-critical-metal-exports-plummet-trade-war-curbs-take-effect.

73. Research and Markets, "EV Battery Recycling Industry Research 2024–2035: Market to Grow at a CAGR of 40.9% with Contemporary Amperex Technology, GEM, Umicore, Glencore, and Fortum Dominating," *GlobeNewswire News Room*, May 7, 2025, https://www.globenewswire.com/news-release/2025/05/07/3076256/28124/en/EV-Battery-Recycling-Industry-Research-2024-2035-Market-to-Grow-at-a-CAGR-of-40-9-with-Contemporary-Amperex-Technology-GEM-Umicore-Glencore-and-Fortum-Dominating.html.

STANFORD EXPERT CONTRIBUTORS

Dr. Steven E. Koonin
SETR Faculty Council and Edward Teller Senior Fellow, Hoover Institution



05

MATERIALS SCIENCE

KEY TAKEAWAYS

- Materials science is a foundational technology that underlies advances in many other fields, including robotics, space, energy, and synthetic biology.
- The field will exploit artificial intelligence as another promising tool to predict new materials with new properties and to identify novel uses for known materials.
- Future progress in materials science requires new funding mechanisms and access to additional computational power to more effectively transition from innovation to implementation.

Overview

From semiconductors in computer chips to plastics in everyday objects, materials are everywhere. Knowing how to synthesize and process them, as well as understanding their structure and properties, has helped to shape the world around us. Materials science contributes to the development of stronger, lighter, and more flexible materials that improve everything from battery electrodes to medical implants and from automobiles to spacecraft.

It is a wide field. At Stanford University, for example, faculty working on materials science research programs are found in many departments, including materials science and engineering, chemical engineering, electrical engineering, bioengineering, chemistry, and physics.

Broadly speaking, materials science research focuses on four major areas:

- **Synthesis of materials** Understanding how materials can be created and assembled from the atomic to macroscopic scale
- **Characterization of materials** Determining their structure and properties, such as conductivity, chemical reactivity, and elasticity
- **Modeling and computational analysis** Studying how materials are formed and how they adapt in specific situations
- **Manufacturing and scaling** Assessing how materials can be produced and scaled for industrial applications

Basics of Materials Science

All materials are composed of atoms. The periodic table of the elements (figure 5.1) lists all the known types of atoms. Certain ones can be combined with others into molecules that have vastly different properties than the individual atoms involved. For example, table salt consists of sodium and chlorine, which are elements. Sodium burns on contact with water, and chlorine is a poisonous gas, yet the table salt we consume every day is a completely different substance.

The periodic table contains ninety-two naturally occurring elements alongside twenty-six laboratory-synthesized ones (with the latter having an atomic number larger than 92). Elements positioned within the same column exhibit roughly similar properties,

FIGURE 5.1 The periodic table of the elements

Group

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1 H																	2 He
2	3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
3	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
4	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
5	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
6	55 Cs	56 Ba	* 71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
7	87 Fr	88 Ra	** 103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Nh	114 Fl	115 Mc	116 Lv	117 Ts	118 Og
			* 57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb		
			** 89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No		

Source: Adapted from Wikimedia Commons, CC BY-SA 4.0

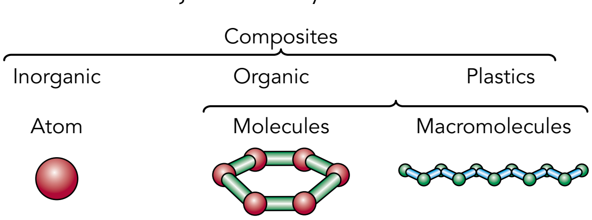
enabling researchers to extrapolate findings from one element to others within its group.

Atoms can be arranged spatially in various ways. A crystal, for example, is the result of arranging atoms in a periodically repeating lattice. The silicon wafer at the heart of the semiconductor industry is one such crystal; more precisely, it's a slice of a single silicon crystal.

Many elements can be combined with one another, generating an extensive array of potential compounds from which materials scientists must identify those with practical applications. This effort is increasingly supported by machine learning (ML) techniques that predict material properties with enough accuracy to expedite identification of promising candidate materials.

Molecules, which are composed of atoms, can, in turn, be linked together into structures called macromolecules (see figure 5.2). These can occur naturally, as is the case for proteins, DNAs, and cellulose, or they can be synthesized artificially and used to create things such as polymers. The polymer chains in plastics dictate the material's properties. If the polymer chains can deform and slip past each other, then the material will be flexible and malleable. The harder it is for the polymer chains to move, the more rigid that material will be. Research on new polymer structures can be used to develop plastics that are easier to recycle or have advantageous mechanical properties while weighing less than metals.

FIGURE 5.2 Objects of study in materials science



Key Developments

Some interesting present-day applications of materials science are discussed below.

Flexible Electronics

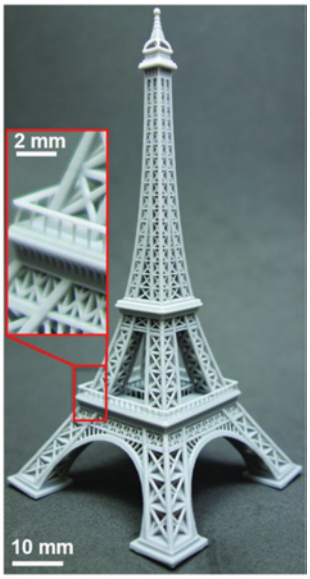
Flexible or stretchable electronics involves the creation of electrical devices that can bend, stretch, and deform without compromising their performance. Such electronics can be used as wearable, skinlike devices. For example, "electronic skin," or e-skin, can conform to real skin and sense things such as temperature and pressure, as well as encode these into electrical signals.[1] A "smart bandage" with integrated sensors to monitor wound conditions and with electrical stimulation can accelerate the time needed to heal chronic wounds by 25 percent.[2]

Additive Manufacturing

One of the most promising advances in materials processing over the past fifteen years is additive manufacturing, colloquially known as 3-D printing. The technology comes in different forms. For instance, a method known as continuous liquid interface production (CLIP) uses directed ultraviolet (UV) light to form structures from a polymer resin.[3] (See figure 5.3.) A key aspect of CLIP is its use of an oxygen-permeable window placed above a UV light projector that prevents the resin from curing in unwanted places.

Especially at high speeds, 3-D printing struggles with producing small features. The 3-D printing process requires several components to perform in concert, including the material resin, the light source, and the build platform where an object is printed. That is technically challenging, but by printing on a tensioned film made from polyethylene terephthalate that's fed through a CLIP printer, it's possible to 3-D print very small particles at a pace of one million a day from a single machine.[4]

FIGURE 5.3 A CLIP-based 3-D printer created a miniature print of the Eiffel Tower



Source: Carbon Inc. / John Tumbleston

Nanotechnology

Nanotechnology is a large subfield of materials science. Size has a profound impact on the properties of a material. Figure 5.4 compares the length of a water molecule (below a nanometer [nm]), a human hair (roughly 10^5 nm), and a human eyeball (at 10^7 nm). A structure is typically referred to as nanoscale if at least one of its dimensions is in the 1 to 100 nm range.

In the past thirty-five years, nanoscience and nanotechnology have attracted enormous interest because the properties of nanoscale materials—including their electronic, optical, magnetic, thermal, and mechanical properties—are often very different from the same material in bulk form.[5] Nanomaterials are classified based on how many of their dimensions are nanoscale:

- **Nanoparticles** have zero dimensions larger than 100 nm.
- **Nanowires** (or nanorods) have one dimension larger than 100 nm (i.e., two dimensions are below 100 nm).
- **Nanosheets** have two dimensions larger than 100 nm (i.e., one dimension is less than 100 nm).
- **Bulk** materials have all three dimensions larger than 100 nm (i.e., no dimensions are less than 100 nm).

FIGURE 5.4 The size of nanoscale objects

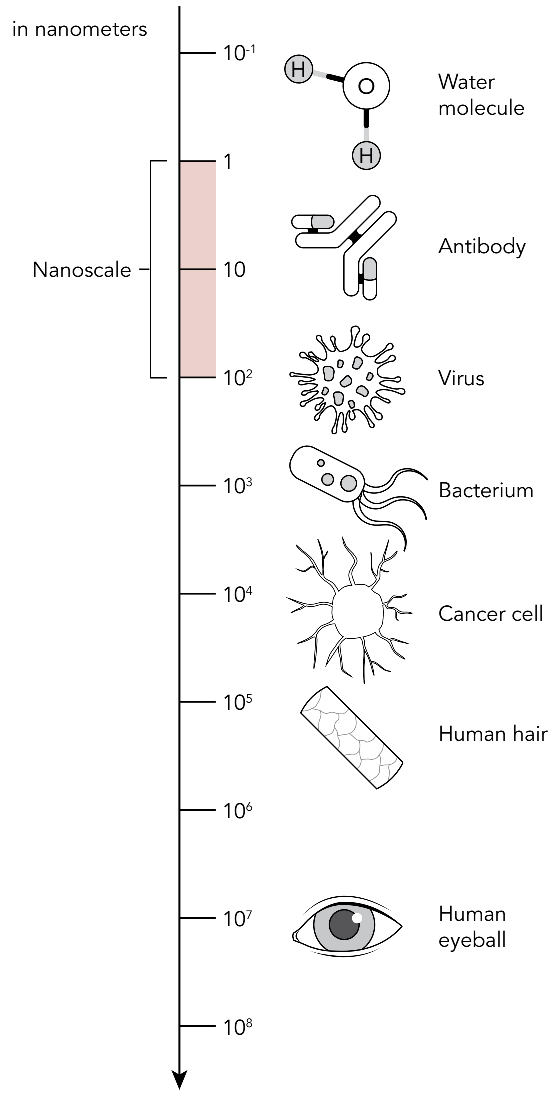
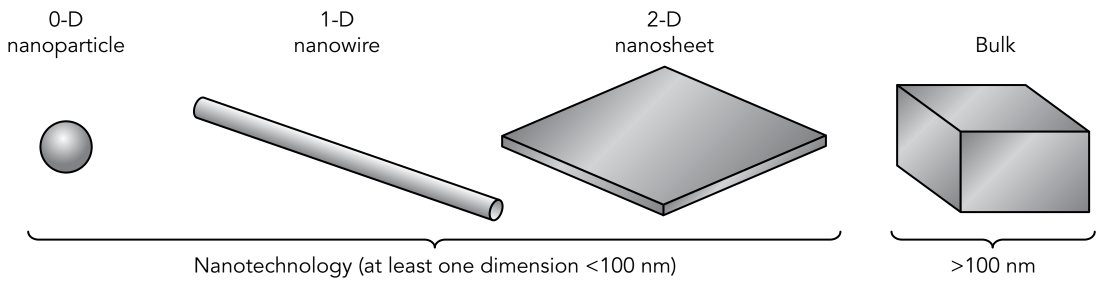


FIGURE 5.5 Dimensionality of nanomaterials



See figure 5.5 for an illustration.

The unique properties of nanomaterials have enabled breakthrough applications across numerous fields, from medicine to electronics. These applications demonstrate how manipulating matter at the nanoscale can solve complex technological challenges and create entirely new possibilities. Some current uses of the technology include the following:

- **Quantum dots** These are metallic, carbonaceous, or semiconductor spherical nanocrystals less than 10 nm in size that emit bright monochromatic light in response to excitation by a light source with a higher energy.[6] Their many applications include being used in medical imaging, as fluorescent markers for biological structures, and in the energy sector, where they enable solar cells to capture more of the solar spectrum.

- **Vaccine stabilization** Vaccines can be encapsulated in lipid nanoparticles (very tiny spheres), making it easier to transport them inside the body and preventing immediate degradation of their contents.[7] This is especially useful for mRNA vaccines, such as the ones developed for COVID-19.

- **Two-dimensional (2-D) semiconductors, graphene, carbon nanotubes, and nanoscale materials** 2-D semiconductors are semiconductors with atomic-scale thickness. These are at the forefront of the next generation of high-tech electronic devices. Active research efforts are designing new methods to integrate 2-D or carbon nanotube semiconductors into electronics that are currently silicon based to improve their energy efficiency and heat management.[8]

Electrochemistry

Electrochemistry studies how electrical energy and chemical reactions interact through electron transfer, typically at electrode-electrolyte interfaces. Electrochemical devices can generate electrical energy through a spontaneous chemical reaction—batteries are a typical example of this—or they can use electrical energy to drive a chemical reaction. Electrocatalytic platforms used to produce hydrogen are a good example of the latter category of device. Materials scientists play a key role in electrochemistry by, for example, developing degradation-resistant battery electrodes and discovering efficient nanocatalysts that drive chemical reactions with electricity.

Batteries

Battery technology has become critical for global energy storage and was recognized by the 2019 Nobel Prize in Chemistry. Current research in materials science addresses three key challenges related to the field: developing better materials for longer-lasting batteries, creating safer alternatives to flammable components, and finding cost-effective substitutes for expensive raw materials. The main obstacles to progress remain achieving higher energy storage levels and faster charging speeds and reducing manufacturing costs—all while ensuring safety and reliability.

Key developments include the following:

- **Nanotechnology in batteries** Using silicon nanowires as battery anodes allows Li-ion batteries to achieve ten times greater energy capacity and maintain stable performance over time. This is because these nanowires can handle large volume changes during charging without breaking apart, overcoming the main limitation of silicon in traditional battery designs.[9]

- **Improved battery cycling** An ML analysis of 186 batteries revealed that faster charging during battery manufacture actually increases battery lifespan by 50 to 70 percent, contradicting traditional slow-charging methods.[10]

- **Solid-state batteries** These replace flammable liquids with ceramics for better safety. However, metal deposits can still penetrate the ceramics, so success depends on preventing manufacturing defects in them.[11]

- **Polymer coatings for safety** Using such coatings with lithium-metal batteries improves their safety. It also helps them achieve over 99.5 percent efficiency and enables them to carry significantly more electrical energy per kilogram than current lithium-ion technology.[12]

- **Battery electrolyte design** ML helped researchers discover that lower oxygen content in electrolyte solvents leads to better cycling (the repeated process of discharging followed by charging) in lithium-metal batteries.[13]

- **Sodium-ion batteries** Unlike lithium-based batteries, which rely on volatile lithium supplies, sodium-ion batteries use abundant materials and could reach cost parity with lithium ones by 2030. Although they may not outperform lithium-based batteries, sodium-ion batteries will enhance supply chain security if they become more widely available.[14]

Electrocatalysis

Electrocatalysis involves using catalysts to accelerate electrochemical reactions. It is essential in processes like water splitting (definition below), fuel cells, and carbon dioxide (CO_2) recycling. Nanomaterials are particularly well suited as electrocatalysts.[15] This is because of their high surface-to-volume ratio, which means many more active catalytic surface areas can participate in a reaction than would be the case for the same material in bulk.

Water splitting uses electricity to convert water into its constituent parts of hydrogen and oxygen. Platinum and other metal nanoparticles are currently used as electrocatalysts to reduce water to hydrogen gas,[16] although scientists are exploring less expensive replacements for them.[17] This process enables renewable energy storage in the form of hydrogen, which can then be consumed in fuel cells to provide on-demand electricity by combusting the hydrogen when needed.

CO_2 electrocatalysis (also known as CO_2 reduction) is a process that uses electricity to convert CO_2 into valuable products such as synthetic fuels, chemicals like methanol and ethylene, and precursors for plastics. It achieves this through the use of specialized catalysts in an electrochemical cell.[18] By transforming CO_2 into essential commodities, this technology provides a route for reducing it in the atmosphere and for storing electricity in chemical form.

> **One of the foremost challenges of materials science as a discipline is the vast number of possible materials and material combinations that can be used and the associated time and cost involved in synthesis and characterization.**

Biosensing

Electrochemistry enables the detection of biological molecules—such as metabolites, hormones, and therapeutic agents—through the use of electrical signals. In electrochemical biosensors, a special biological component like an enzyme, antibody, single strand of DNA, or other material is placed on an electrode and reacts specifically with the substance the sensor is trying to detect. This interaction generates or alters an electrochemical signal (e.g., an electrical current) that is measured by the electrode. The high sensitivity, low cost, and portability of electrochemical biosensors, such as wearable glucose sensors for managing diabetes, make them ideal for medical diagnostics, environmental monitoring, and pathogen detection.

The Application of Artificial Intelligence in Materials Science

One of the foremost challenges of materials science as a discipline is the vast number of possible materials and material combinations that can be used and the associated time and cost involved in synthesis and characterization. ML offers a solution by recognizing patterns in existing data that can help models make predictions about new materials and their properties.[19] While this approach has been successful with relatively simple materials, complex ones remain difficult to predict.

To truly understand and forecast the properties of materials, more accurate, comprehensive, and tailored databases are needed. Their many applications could include helping accelerate the development of materials that enable researchers to overcome bottlenecks in chip assembly as semiconductors continue to be miniaturized. Databases such as the Materials Project, led by Lawrence Berkeley National Laboratory, represent the significant effort being made to improve data gathering.[20] However, these efforts are still limited with respect to the range of materials' properties that are covered. Major companies are already using artificial intelligence (AI) to discover new materials. Google DeepMind's Graph Networks project uses neural networks to predict material properties, while IBM, Citrine Informatics, and MaterialsZone combine materials expertise with data science to speed up development and improve product design.

To tackle the problem of limited experimental data, researchers are developing autonomous laboratories that can quickly synthesize and test materials at scale. The A-Lab at University of California–Berkeley (figure 5.6) exemplifies this approach, using robotic arms and AI-guided synthesis alongside automated characterization equipment to discover new materials.[21] However, the scientific community has questioned the accuracy of A-Lab's characterization work and structural analysis, with researchers noting areas for improvement.[22] This emphasizes that computational and integrated approaches still need careful validation and human oversight but hold great promise. Future automated labs may be capable of accurately predicting and creating new

FIGURE 5.6 The A-Lab combines AI-guided synthesis with automated materials characterization

Source: © 2023 The Regents of the University of California, Lawrence Berkeley National Laboratory

materials, reducing the human effort needed to design them.

By combining advanced algorithms with expanded databases and automated experimentation, researchers are working on exciting efforts to explore the vast landscape of possible materials more efficiently than ever before.

Over the Horizon

Enhancing the Human Body Through Materials Science

Advancements in materials science are revolutionizing how we repair, restore, and augment the human body. From supporting regenerative medicine to brain–machine interfaces, engineered materials are enabling technologies once thought to be science fiction. Prosthetics and bionics have made major leaps thanks to high-performance materials. Lightweight composites like carbon fiber and shape memory alloys allow artificial limbs to move naturally and respond dynamically. Electrically conductive and biocompatible polymers help to form soft bioelectronic interfaces that are powering neuroprosthetics.[23] These prosthetics aim to restore and enhance motor and sensory function by connecting directly with the nervous system.

A particularly exciting frontier is brain–machine interfaces, which depend on materials that are both biocompatible and capable of recording or stimulating neural activity. (These interfaces are discussed further in chapter 6, on neuroscience.) Pliant, ultrathin electrodes made from composite materials, such as flexible polymers embedded with 2-D conductors like graphene, carbon nanotubes, and metal nanowires, are designed to conform to the brain's surface without damaging tissue.[24] These materials interface seamlessly with neurons and can detect and decode brain signals with high resolution. This, in turn, enables the execution of computer tasks and the control of robotic limbs, providing a means of communication for paralyzed patients.

While scientists improve brain–machine interfaces for neural activity decoding, the next goal is active neuromodulation, or the ability to control neural activity with an implanted device. Currently, deep-brain neuromodulatory devices are being used to successfully mitigate symptoms of Parkinson's disease. Researchers are also working with them to affect mood and memory.[25] Such advances suggest that our ability to engineer materials is enabling the construction of devices that are fundamentally expanding what the human body and mind can do.

Metamaterials: Programming Physical Properties

Metamaterials are artificially engineered materials with optical or acoustic properties not found in nature. These properties arise from arranging engineered microscopic structural components in particular patterns, with feature sizes smaller than the wavelength of light or sound that is of interest. The internal structure of metamaterials enables extraordinary manipulation of electromagnetic or acoustic waves, including bending them in previously impossible ways, opening new possibilities for controlling light or sound.

In recent years, the field has shifted from proof-of-concept demonstrations to various applications.[26] These include:

- **Invisibility cloaks** Bending light around objects to render them nearly invisible to the naked eye
- **Superlenses** Imaging with resolution beyond the diffraction limit in microscopes and medical imaging
- **Advanced radio antennas** Creating compact, efficient antennas with improved signal strength for communications
- **Seismic protection** Shielding buildings from earthquakes by redirecting seismic waves
- **Acoustic control** Improved materials for sound insulation, muffling vibrations, or sound filtering in medical and industrial settings

Advances in manufacturing and fabrication have been the main reason for the implementation of metamaterials in the physical world. Advances in laser processing and multi-material processing, among other things, have made complex metamaterial designs commercially viable at scale.

In the future, new applications in telecommunications, biomedicine, and energy are expected.[27] Reconfigurable intelligent antenna surfaces for 6G communications show promise for future growth. Biomedical applications include metamaterial-enhanced wireless power for implants and multi-disease diagnosis with high-frequency biosensors. Metamaterials are also well positioned as a foundational technology for next-generation systems spanning from quantum computing to autonomous sensing.

Policy Issues

Research Infrastructure

Today's materials science research infrastructure does not adequately support the transition from research to real-world applications at scale. Such transitions generally require construction of a small-scale pilot project to demonstrate the feasibility of potential large-scale manufacturing. At this point, the technology is too mature to qualify for most research funding—because basic science does not address issues related to scaling up—but not mature enough to be commercialized by actual companies.

Neither government funders nor venture capital investors are particularly enthusiastic about financing

Historically, the United States has led the world in nanotechnology, but the gap between it and China has narrowed.

pilot projects given the significant up-front investment needed. Therefore, different forms of funding are required to bridge the gap between bench-scale research and company-level investment. Although the US government does occasionally support pilot projects, additional support in this area could also establish national rapid prototyping centers where academic researchers can find the help and tools necessary to build prototypes and pilot plants for their technology.

Today's research processes are also ill-suited to rapid transitions to real-world applications. Such processes emphasize sequential steps. The standard process has been to characterize a material and then proceed to a simple demonstration of how it might be used. Today, addressing big societal challenges calls for a more scalable, system-level approach that involves extensive rapid prototyping and fast, reliable demonstrations to provide feedback on the potential value of specific materials and to fill in knowledge gaps.

Current research arrangements make this difficult. For example, in collaborations with a medical school, it is often necessary to bring almost-finished products to clinical tests to validate the true impact of a new medical device that is using innovative materials. There is typically a window of thirty minutes or less in which to place a device on a patient and gather data. This means that any malfunction, such as a sudden equipment failure or a loose wire, can jeopardize an entire experiment and potentially halt future patient interactions. Lab-assembled devices may not meet this standard of reliability, even if they do demonstrate the value of the underlying science.

Regulation of Products Incorporating (Nano)Materials and Environmental Concerns

As with other areas of technology, materials science faces concerns about the appropriate regulatory balance between the need to ensure public safety and the imperative to innovate quickly and leapfrog possible competitors. In the biomedical space, the US Food and Drug Administration (FDA) created a Nanotechnology Regulatory Science Research Plan in 2013.[28] Today, FDA regulation and review of nanotechnology is governed by Executive Order 13563.[29] Outside of biomedicine, regulation of and infrastructure for nanomaterials research from the government side is based largely in the agencies involved in the National Nanotechnology Initiative. These include the US Department of Energy, the National Cancer Institute, the National Institutes of Health, the National Institute of Standards and Technology in the Department of Commerce, and the National Science Foundation.

Nanoparticles raise particular concerns because their small size may enable them to pass through various biological borders, such as cell membranes or the blood–brain barrier, potentially harming biological systems. Nanoscale particles inhaled into the lungs, for example, may lodge there permanently, causing severe health outcomes, including pulmonary inflammation, lung cancer, and penetration into the brain and skin.[30] Nanoparticles may result from air pollution and even doing laundry. However, the most common source of nanoparticles is through combustion linked to the mechanical abrasion of common objects such as cookware and car tires. By comparison, laboratory-engineered nanoparticles are an insignificant source of nanoparticle contamination today.

Finally, end-of-life considerations that take into account environmental sustainability and resource conservation are inherently a part of developing and distributing new materials. This is especially important for plastics and materials containing per- and polyfluoroalkyl substances (PFAS), which pose significant environmental and health risks. Material developers can incorporate recyclability into their design processes. For PFAS and other persistent chemicals, the US Environmental Protection Agency strategic road map of October 2021 provides guidance for their use and disposal, and calls for research into safe alternatives and effective degradation methods.[31]

The Mineral Supply Chain

Sourcing minerals involves a complex, global network that encompasses the mining, processing, and distribution of critical raw elements such as lithium, cobalt, nickel, copper, and rare-earth elements (e.g., neodymium and yttrium). These minerals are in high demand for technologies like batteries, renewable energy systems, and consumer and defense electronics.

The mining of these minerals is often concentrated in a few sites that are controlled by countries that may not be aligned with the United States. For example, the Democratic Republic of the Congo exports 74 percent of the world's cobalt from mines that exploit slave and child labor and are operated by Chinese companies.[32] Additionally, China accounts for over 90 percent of the refining capacity of rare-earth elements, which are critical for the manufacturing of magnets used in modern electronics and other industries.[33] This means China can unilaterally squeeze the supply of key minerals and negatively impact electronics manufacturing in the United States and other countries.

It is imperative that America's supply chain for minerals becomes more secure. Solutions include diversifying sources, investing in recycling and the reuse of materials that would otherwise have to be mined, and developing alternative materials. Strengthening domestic processing capabilities and fostering international partnerships are also critical strategies to enhance security, transparency, and resilience in the mineral supply chain.

Foreign Collaboration and Competition

Historically, the United States has led the world in nanotechnology, but the gap between it and China has narrowed. For example, in 2024, China's output of publications in nanotechnology was about 5.3 times higher than that of the United States.[34]

As great power competition intensifies, many researchers are concerned that fundamental research in the United States could now be subject to export controls. Policy ambiguity can inadvertently hinder innovation by creating obstacles for foreign researchers wishing to contribute to work in America and by deterring international collaborations with allies and partners who are important for advancing the field.

It is essential that scholars can collaborate broadly on fundamental research at an international level so that the exchange of ideas and perspectives can foster new ways of thinking that increase the likelihood and speed of technological breakthroughs. In nanomaterials, for example, researchers in South Korea are making significant strides with biomedical applications and applications for consumer electronics. There is an urgent need for clarification of these policies, particularly those delineating fundamental research and export-controlled research.

Infrastructure for ML-Assisted Materials Science

The United States benefits from having some of the world's largest supercomputing resources, which are essential not only for ML but also for developing extensive databases. However, better access to computing power is necessary for researchers in materials science to generate and analyze databases effectively. Greater access to data, including to databases that might not always be openly available to academics, is also needed.

One additional area where policymakers could have a significant impact is in bridging the gap between the scientific community and makers of computational hardware. Frequent changes in computing architectures can lead to a loss of productivity for researchers because code must be constantly updated. Improved collaboration with hardware manufacturers and other providers of computing resources could ensure scientific needs are better aligned with advances in computing technology, enhancing overall research efficiency.

NOTES

1. Weichen Wang, Yuanwen Jiang, Donglai Zhong, et al., "Neuromorphic Sensorimotor Loop Embodied by Monolithically Integrated, Low-Voltage, Soft E-Skin," *Science* 380, no. 6646 (2023): 735–42, https://doi.org/10.1126/science.ade0086.

2. Yuanwen Jiang, Artem A. Trotsyuk, Simiao Niu, et al., "Wireless, Closed-Loop, Smart Bandage with Integrated Sensors and Stimulators for Advanced Wound Care and Accelerated Healing," *Nature Biotechnology* 41 (2023): 652–62, https://doi.org/10.1038/s41587-022-01528-3.

3. John R. Tumbleston, David Shirvanyants, Nikita Ermoshkin, et al., "Continuous Liquid Interface Production of 3D Objects," *Science* 347, no. 6228 (2015): 1349–52, https://doi.org/10.1126/science.aaa2397; Kaiwen Hsiao, Brian J. Lee, Tim Samuelsen, et al., "Single-Digit-Micrometer-Resolution Continuous Liquid Interface Production," *Science Advances* 8, no. 46 (2022), https://doi.org/10.1126/sciadv.abq2846.

4. For more on polyethylene terephthalate, see Polymershapes, "PET Plastic Film," accessed October 13, 2023, https://www.polymershapes.com/pet-plastic-film/; Jason M. Kronenfeld, Lukas Rother, Max A. Saccone, Maria T. Dulay, and Joseph M. DeSimone, "Roll-to-Roll, High-Resolution 3D Printing of Shape-Specific Particles," *Nature* 627 (March 2024): 306–12, https://doi.org/10.1038/s41586-024-07061-4.

5. Hui Pan and Yuan Ping Feng, "Semiconductor Nanowires and Nanotubes: Effects of Size and Surface-to-Volume Ratio," *ACS Nano* 2, no. 11 (2008): 2410–14, https://doi.org/10.1021/nn8004872; Anna C. Balazs, Todd Emrick, and Thomas P. Russell, "Nanoparticle Polymer Composites: Where Two Small Worlds Meet," *Science* 314, no. 5802 (2006): 1107–10, https://doi.org/10.1126/science.1130557.

6. A. P. Alivisatos, "Semiconductor Clusters, Nanocrystals, and Quantum Dots," *Science* 271, no. 5251 (1996): 933–37, https://doi.org/10.1126/science.271.5251.933.

7. Norbert Pardi, Michael J. Hogan, Frederick W. Porter, and Drew Weissman, "mRNA Vaccines—A New Era in Vaccinology," *Nature Reviews Drug Discovery* 17 (2018): 261–79, https://doi.org/10.1038/nrd.2017.243.

8. Weisheng Li, Xiaoshu Gong, Zhihao Yu, et al., "Approaching the Quantum Limit in Two-Dimensional Semiconductor Contacts," *Nature* 613 (2023): 274–79, https://doi.org/10.1038/s41586-022-05431-4; Eric Pop, "Energy Dissipation and Transport in Nanoscale Devices," *Nano Research* 3 (2010): 147–69, https://doi.org/10.1007/s12274-010-1019-z.

9. Candace K. Chan, Hailin Peng, Gao Liu, et al., "High-Performance Lithium Battery Anodes Using Silicon Nanowires," *Nature Nanotechnology* 3 (December 2008): 31–35, https://doi.org/10.1038/nnano.2007.411.

10. Xiao Cui, Stephen Dongmin Kang, Sunny Wang, et al., "Data-Driven Analysis of Battery Formation Reveals the Role of Electrode Utilization in Extending Cycle Life," *Joule* 8, no. 11 (2024): 3072–87, https://doi.org/10.1016/j.joule.2024.07.024.

11. B. Zhang, B. Yuan, X. Yan, et al., "Atomic Mechanism of Lithium Dendrite Penetration in Solid Electrolytes," *Nature Communications* 16, 1906 (2025), https://doi.org/10.1038/s41467-025-57259-x.

12. Zhuojun Huang, Hao Lyu, Louisa C. Greenburg, Yi Cui, and Zhenan Bao,"Stabilizing Lithium-Metal Electrodes with Polymer Coatings," *Nature Energy* 10, no. 7 (2025): 811–23, https://doi.org/10.1038/s41560-025-01767-z.

13. Sang Cheol Kim, Solomon T. Oyakhire, Constantine Athanitis, et al., "Data-Driven Electrolyte Design for Lithium Metal Anodes," *Proceedings of the National Academy of Sciences* 120, no. 10 (2023): e2214357120, https://doi.org/10.1073/pnas.2214357120.

14. Adrian Yao, Sally M. Benson, and William C. Chueh, "Critically Assessing Sodium-Ion Technology Roadmaps and Scenarios for Techno-Economic Competitiveness against Lithium-Ion Batteries," *Nature Energy* 10, no. 3 (2025): 404–16, https://doi.org/10.1038/s41560-024-01701-9.

15. U. P. M. Ashik, Anchu Viswan, Shinji Kudo, and Jun-ichiro Hayachi, "Nanomaterials as Catalysts," chap. 3 in *Applications of Nanomaterials: Advances and Key Technologies*, eds. Sneha Mohan Bhagyaraj, Oluwatobi Samuel Oluwafemi, Nandakumar Kalarikkal, and Sabu Thomas (Elsevier, 2018), 4582, https://doi.org/10.1016/B978-0-08-101971-9.00003-X.

16. Thomas F. Jaramillo, Kristina P. Jørgensen, Jacob Bonde, Jane H. Nielsen, Sebastian Horch, and Ib Chorkendorff, "Identification of Active Edge Sites for Electrochemical H_2 Evolution from MoS*2* Nanocatalysts," *Science* 317, no. 5834 (2007): 100–102, https://doi.org/10.1126/science.1141483.

17. Zhi Wei Seh, Jakob Kibsgaard, Colin F. Dickens, Ib Chorkendorff, Jens K. Nørskov, and Thomas F. Jaramillo, "Combining Theory and Experiment in Electrocatalysis: Insights Into Materials Design," *Science* 355, no. 6321 (2017), https://doi.org/10.1126/science.aad4998.

18. Xiaolong Zhang, Si-Xuan Guo, Karl A. Gandionco, Alan M. Bond, and Jie Zhang, "Electrocatalytic Carbon Dioxide Reduction: From Fundamental Principles to Catalyst Design," *Materials Today Advances* 7 (2020), https://www.sciencedirect.com/science/article/pii/S2590049820300217.

19. Steven G. Louie, Yang-Hao Chan, Felipe H. da Jornada, Zhenglu Li, and Diana Y. Qiu, "Discovering and Understanding Materials Through Computation," *Nature Materials* 20 (2021): 728–35, https://doi.org/10.1038/s41563-021-01015-1.

20. The Materials Project, "The Materials Project," accessed September 18, 2024, https://next-gen.materialsproject.org.

21. Nathan J. Szymanski, Bernardus Rendy, Yuxing Fei, et al., "An Autonomous Laboratory for the Accelerated Synthesis of Novel Materials," *Nature* 624 (November 2023): 86–91, https://doi.org/10.1038/s41586-023-06734-w.

22. Josh Leeman, Yuhan Liu, Joseph Stiles, et al., "Challenges in High-Throughput Inorganic Material Prediction and Autonomous Synthesis," preprint, ChemRxiv, submitted January 7, 2024, https://doi.org/10.26434/chemrxiv-2024-5p9j4.

23. Jing Li et al., "PEDOT:PSS-Based Bioelectronics for Brain Monitoring and Modulation," *Microsystems & Nanoengineering* 11 (2025): article 87, https://www.nature.com/articles/s41378-025-00948-w.

24. Dimitris Boufidis et al., "Bio-Inspired Electronics: Soft, Biohybrid, and 'Living' Neural Interfaces," *Nature Communications* 16 (2025): article 1861, https://www.nature.com/articles/s41467-025-57016-0.

25. Andres M. Lozano et al., "Deep Brain Stimulation: Current Challenges and Future Directions," *Nature Reviews Neurology* 15 (2019): 148–60. https://www.nature.com/articles/s41582-018-0128-2.

26. Rakesh Kumar, Manoj Kumar, Jasgurpreet Singh Chohan, and Santosh Kumar, "Overview on Metamaterial: History, Types and Applications," *Materials Today: Proceedings* 56 (no. 5): 3016–24, https://doi.org/10.1016/j.matpr.2021.11.423.

27. J. U. Surjadi and C. M. Portela, "Enabling Three-Dimensional Architected Materials Across Length Scales and Timescales," *Nat. Mater* 24 (2025): 493–505, https://doi.org/10.1038/s41563-025-02119-8.

28. US Food and Drug Administration, "2013 Nanotechnology Regulatory Science Research Plan," last modified March 19, 2018, https://www.fda.gov/science-research/nanotechnology-programs-fda/2013-nanotechnology-regulatory-science-research-plan.

29. The White House, "Executive Order 13563–Improving Regulation and Regulatory Review," Office of the Press Secretary, January 18, 2011, https://obamawhitehouse.archives.gov/the-press-office/2011/01/18/executive-order-13563-improving-regulation-and-regulatory-review.

30. Paresh Chandra Ray, Hongtao Yu, and Peter P. Fu, "Toxicity and Environmental Risks of Nanomaterials: Challenges and Future Needs," *Journal of Environmental Science and Health*, Part C 27 (February 2009): 1–35, https://doi.org/10.1080/10590500802708267.

31. EPA Council on PFAS, "PFAS Strategic Roadmap: EPA's Commitment to Action 2021–2024," US Environmental Protection Agency, October 18, 2021, epa.gov/system/files/documents/2021-10/pfas-roadmap_final-508.pdf.

32. Terry Gross, "Red Cobalt in Congo: How Mining Fuels Human Suffering," NPR, February 1, 2023, https://www.npr.org/sections/goatsandsoda/2023/02/01/1152893248/red-cobalt-congo-drc-mining-siddharth-kara.

33. Gracelin Baskaran and Meredith Schwartz, "The Consequences of China's New Rare Earths Export Restrictions," Critical Questions, Center for Strategic and International Studies, April 14, 2025, https://www.csis.org/analysis/consequences-chinas-new-rare-earths-export-restrictions.

34. "Top 10 Countries by Nanotechnology Publications in 2024," accessed August 19, 2025, https://statnano.com/news/print/74558/Top-10-Countries-by-Nanotechnology-Publications-in-2024.

STANFORD EXPERT CONTRIBUTORS

Dr. Zhenan Bao
SETR Faculty Council, K. K. Lee Professor of Chemical Engineering, and Professor, by courtesy, of Materials Science and Engineering and of Chemistry

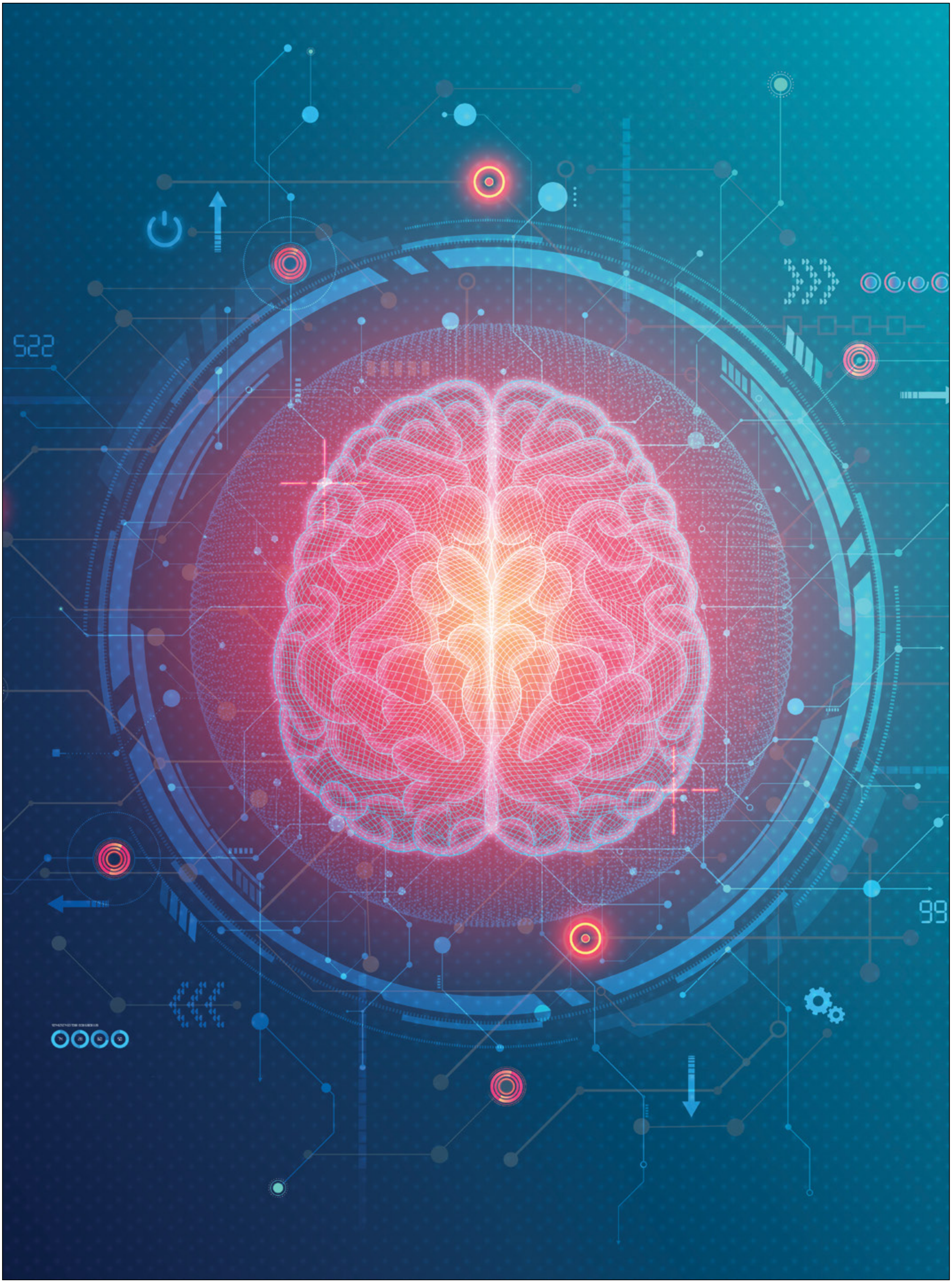
Dr. Felipe da Jornada
Assistant Professor of Materials Science and Engineering

Dr. Matthew Kanan
Professor of Chemistry and Senior Fellow at the Precourt Institute for Energy

Dr. Andrew Spakowitz
Tang Family Foundation Chair of the Department of Chemical Engineering and Professor of Chemical Engineering and of Materials Science and Engineering

Dr. Stefano Cestellos-Blanco
SETR Fellow and Postdoctoral Scholar in Chemical Engineering

Dr. Lukas Michalek
SETR Fellow and Postdoctoral Scholar in Chemical Engineering


522
99

06

NEUROSCIENCE

KEY TAKEAWAYS

- Advances in human genetics and experimental neuroscience, along with computing and neuroscience theory, have led to some progress in several areas, including understanding and treating addiction and neurodegenerative diseases and designing brain–machine interfaces for restoring vision.

- American leadership is essential for establishing and upholding global norms about ethics and human subjects research in neuroscience, but this leadership is slipping with decreased strategic planning and increased foreign investments in the field.

- Popular interest in neuroscience vastly exceeds the current scientific understanding of the brain, giving rise to overhyped claims in the public domain that revolutionary advances are just around the corner.

Overview

Neuroscience is a multidisciplinary field of study that focuses on the components, functions, and dysfunctions of the brain and our nervous system at every level. It reaches from the earliest stages of embryonic development to dysfunctions and degeneration later in life, and its study spans from the individual molecules that shape the functions of a neuron to the complex system dynamics that constitute our thoughts and dictate our behaviors.

The human brain consumes 20 to 25 percent of the body's energy even though it constitutes only a small percentage of a human's body weight, a fact that underscores its outsize importance.[1] The power of the human brain is what has allowed us to become the dominant species on Earth without being the fastest, strongest, or biggest.

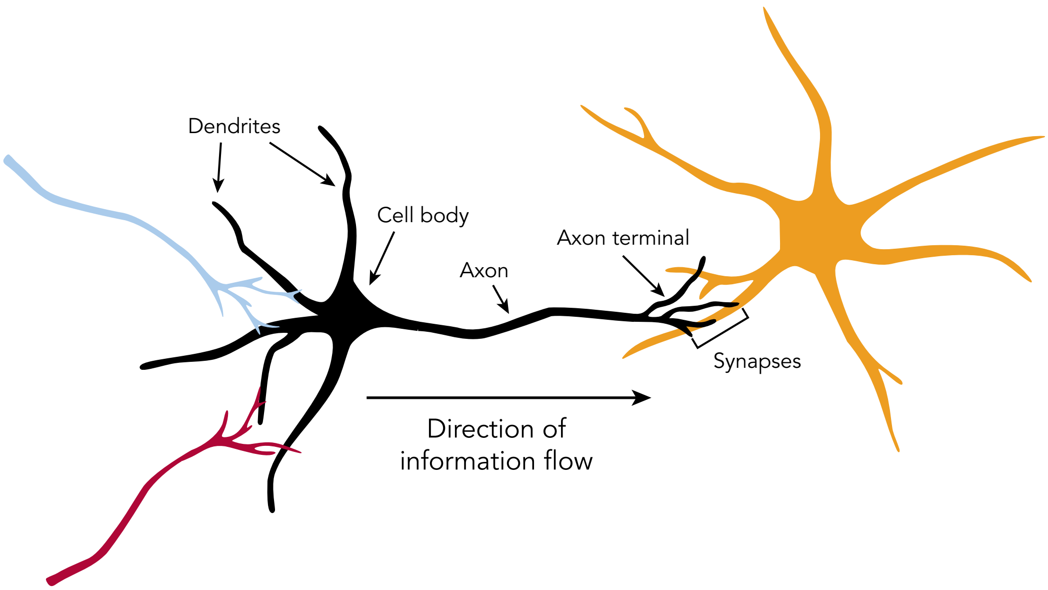
The brain is unfathomably complex, containing approximately eighty-six billion neurons[2]—nerve cells that sense the physical world, transmit information to the brain, process information, and send information from the brain to other parts of the body. A single neuron can make thousands or tens of thousands of connections to other neurons. These connections are called synapses (see figure 6.1).

All of our consciousness and behavior, from the action of stabbing a potato with a fork to contemplating the mysteries of the universe, is underpinned by which neurons connect with one another, the neurotransmitter/receptor pairs involved, the strength of the connections, and the electrical properties of the neurons—as well as by how these various features change over time.

Neurons and synapses function in many ways that are similar to electrical circuits. Indeed, the exploration of the electrical properties of neurons came directly from the same technologies, theories, and equations developed for harnessing electricity. Many pioneering neuroscientists started as electrical engineers and physicists. Just as electrical connectors create a path for electricity to flow through a circuit, neural circuits can be defined by the parallel and recurrent connections between neurons that occur to compute a specific function, such as deciding to move a limb or visually identifying an object. Neurons can also communicate with each other using hormone-like signaling, which is relatively slow but longer lasting compared to fast-acting electric signals. These types of communications underlie mood and behavior states such as sleep/awake and hunger/satiety.

Complete understanding of what each neuron is doing at any given time is currently impossible. Even for a mouse brain, which is much simpler than a human brain, it is still a tremendous effort to characterize individual brain regions despite the availability of powerful techniques that allow us to identify activity in individual neurons or to noninvasively tag cells to respond to light signals.

FIGURE 6.1 Structure of a neuron



Headlines about mind-reading chip implants are . . . still more the realm of science fiction.

Over the past several years, however, it has become clear that individual neurons are almost never responsible for any given behavior or computation; instead, they act in parallel, duplicating some functions and combining to determine thoughts and actions. This neural redundancy makes it easier to infer what is going on in the brain more broadly.

A particular brain region can be considered like a magnificent choir of a thousand voices. Sampling just 1 percent of the singers can provide a pretty good idea of the music the overall choir is producing at any given time. Researchers already have the ability to record from thousands of neurons at a time. This provides useful insight into how a brain functions, even if we don't understand in detail what the other 99 percent of its neurons are doing.

It is important to keep in mind that the pace of neuroscientific discovery is slow and limited by the biological nature and complexity of the nervous system. Year-over-year advances tend to be incremental. Furthermore, the brain's complexity often prevents researchers from fully understanding why even effective treatments for neurological conditions actually work. For example, we know that drugs called selective serotonin reuptake inhibitors block the reabsorption of serotonin into neurons, but neuroscientists do not have a clear explanation for why this helps treat depression. Sometimes, even if a detailed understanding of a treatment's mechanisms is not essential for therapeutic intervention, knowledge of the underlying biology greatly aids in the search for new drugs and therapies.

To gain understanding of how molecules and neurons work, many researchers use simple model organisms like fruit flies and mice to study fundamental questions inexpensively. But the closer research gets toward human application, the more complex, time-consuming, and expensive it becomes. For instance, because neurodegeneration is a slow, progressive disease where day-to-day worsening is minimal, clinical trials often take many years.

Key Developments

This chapter focuses on three research areas in neuroscience that show major promise for concrete applications: brain–machine interfaces (neuroengineering), degeneration and aging (neurohealth), and the science of addiction (neurodiscovery). Most of the economic impacts of neuroscience connect in some way to the healthcare industry and its search for treatments for neurodegenerative disorders (such as Alzheimer's and Parkinson's disease) and neuropsychiatric disorders (addiction, depression, and schizophrenia) and neural prosthesis (brain–machine interfaces to restore limb function and speech).

Neuroengineering and the Development of Brain–Machine Interfaces

A brain–machine interface is a device that maps neural impulses from the brain and translates these signals to computers. The potential applications for mature brain–machine interface technologies are wide-ranging: The augmentation of vision, other senses, and physical mobility; direct mind-to-computer interfacing; and computer-assisted memory recall and cognition are all within the theoretical realms of possibility. However, headlines about mind-reading chip implants are exaggerated and still more the realm of science fiction. Even with tremendous interest and increasing progress in neuroscience and engineering, the necessary theoretical understanding of how neurocircuits work is still limited to only a few areas of the brain. What's more, the technical problems of safely implanting electrodes have not been solved.

One encouraging example of a brain–machine interface is the recent development of an artificial retina. The retina is the part of the eye that converts light into corresponding electrical signals sent to the brain. People who have certain incurable retinal diseases are blind because the light-detecting cells in their retinas do not work. To restore sight, the Stanford artificial retina project aims to take video images and use electrodes implanted in the eye to simulate the electronic signals in a pattern that a functional retina would normally produce.[3]

The project involves recording spontaneous neural activity to identify cell types and their normal signals, understanding how electrodes activate cells, and stimulating retinal ganglion cells—which collect visual information from photoreceptors in retinas—to represent an image so that this information can be transmitted by the optic nerve to the brain. Solving these technical problems calls for deep knowledge of relevant surgical techniques as well as significant engineering know-how in multiple areas; this includes translating the scientific understanding of the stimulation algorithm used into practical applications, making experimental recordings, and fabricating and packaging the electrode into the device.

The artificial retina project is the most mature brain–machine interface to date in terms of its ability to "read" and "write" information. The retina, a part of the central nervous system, is well suited as an experimental environment because its stimuli (light) is experimentally controllable and can be captured by a digital camera. It is the best-understood neural circuit, and the theory of its function has developed to the point where much of retinal processing can be modeled. Compared to complex cognitive processes like learning and memory—where even the inputs aren't fully understood—the task of reconstructing vision is more achievable, albeit still challenging.

Other brain–machine interfaces are currently being developed, though they are less mature or less ambitious than the artificial retina project. Some of these decode brain activity without controlling a neural signal. For instance, one interface can translate brain activity in areas controlling motor functions into signals that can then be sent to an artificial prosthetic limb. Here, feeding high-dimensional patterns of recorded neural activity into an artificial intelligence (AI) algorithm can make it possible to control an artificial limb without requiring direct control of neural functions—a form of control that remains beyond our current scientific understanding.

These demonstrations hint at the prospect of other brain–machine interfaces in the future, such as computer-assisted memory recall, even if the full suite of potential applications is still unclear. The scope and feasibility of these applications will be determined by advances in neuroscientific theory and by technical solutions to engineering problems such as how to safely and accurately insert probes into deep-layer tissues.

Neurohealth and Neurodegeneration

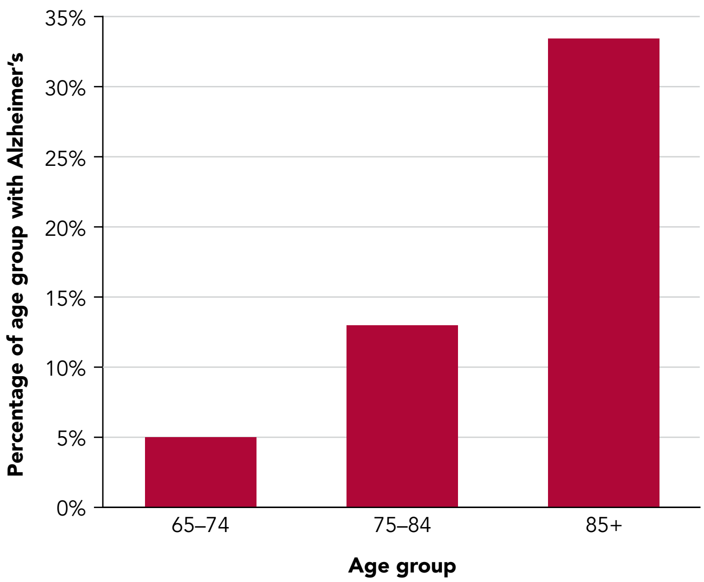
Neurodegeneration is a major challenge as humans live longer. Alzheimer's disease is of particular concern. In the United States alone, the annual cost of treating it is projected to grow from $305 billion in 2020 to $1 trillion by 2050.[4] Diseases like Alzheimer's and Parkinson's surge in frequency with age; while just 5 percent of 65- to 74-year-olds have Alzheimer's, this rises to 33 percent for those over 85 (see figure 6.2).[5] As modern medicine and society enable longer lifespans, the human body and brain remain maladapted to maintaining nervous system function for decades past childbearing age.

Alzheimer's disease is characterized by the accumulation of two different proteins—amyloid beta and tau—into toxic aggregates. Amyloid beta accumulates outside of neurons, induces cellular stress, and in turn may cause tau to build up inside the neurons. As the brain regions where tau accumulates are those most cognitively impacted, a reasonable consensus exists that tau is the more direct cause of the neural death responsible for dementia.

However, despite what is known about neurodegenerative diseases such as Alzheimer's, no drugs can reverse the associated memory loss. Tau remains harder to target therapeutically, and the recently approved drugs target amyloid beta. While these amyloid drugs are very effective in eliminating the amyloid plaques from patient brains, they only modestly slow disease progression. One factor that can improve the outcomes of future clinical trials is to select early-stage patients whose neurons have not yet died, as re-creating memories after neurons have died is not believed to be possible. Efforts are being made to develop techniques for early diagnosis for more effective intervention.

Another form of neurodegeneration results from traumatic brain injury (TBI), which can manifest itself in a range of complex symptoms and pathologies.[6] Traumatic impact to brain systems can affect cognitive and behavioral functions in ways that lead to long-term and severe psychiatric conditions requiring specialized care. This is particularly evident in the current surge of athletic and military brain injuries that exhibit predominantly psychiatric symptoms. A person's past medical and psychiatric records, as well as any coexisting conditions, play a vital role in diagnosis and treatment. TBI offers insights into other neuropsychiatric disorders and can pave the way for innovative concepts in neurodegenerative disease.

FIGURE 6.2 Alzheimer's disease surges in frequency with age



Source: Data from "2023 Alzheimer's Disease Facts and Figures," *Alzheimer's & Dementia* 19, no. 4 (April 2023): 1598–1695, https://doi.org/10.1002/alz.13016

Neurodiscovery and the Science of Addiction

Researchers are working to understand the neural bases of addiction and chronic pain while collaborating with psychiatrists and policymakers to address the opioid epidemic.[7] Estimates of the economic costs of that epidemic range from $100 billion to $1 trillion a year when the loss of potential lifetime earnings of overdose victims is included.[8] Additional economic losses occur due to depletion of the labor force and the billions spent on the criminal justice system and healthcare related to addiction.[9] Beyond economics, there are the significant emotional costs to individuals experiencing addiction and their families and friends. Death also takes its toll: The number of opioid deaths in the United States rose from 21,000 in 2010 to 83,000 in 2022,[10] placing deaths from opioid overdoses at the same level as those caused by diabetes and Alzheimer's.[11] Overdose deaths from opioids fell by 38 percent between the end of 2023 and the end of the following year,[12] but it is unclear if that trend will continue.

Many of the most impactful changes for dealing with the societal problems arising from addiction come from public policy interventions and societal shifts, such as raising taxes on tobacco or changing physicians' prescribing practices for addictive substances such as opioids (see figure 6.3). Nevertheless, neuroscience has a potentially important role to play in addressing addiction. For example, a nonaddictive painkiller drug as effective as current-generation opioids could be transformative.[13]

Another approach is to leverage neuroscience to identify and target brain states that reinforce addiction or make it more likely. Consider the problem of relapse in tackling addiction. Scientists have found that the brain mechanisms leading to an initial opioid addiction differ significantly from those that trigger a relapse. It turns out that opioid receptors are found in neural circuits related to the desire for social interaction. Stanford neuroscientists have recently identified a circuit that is responsible for the onset of aversion to social interactions during recovery.[14] Such an aversion is a significant challenge to recovery because social interactions are often key to helping an individual cope with the vulnerabilities associated with the recovery process. The finding suggests it may be possible to develop drugs that inhibit social aversion during withdrawal, thereby assisting patients in seeking help or companionship from friends, families, recovery programs, and doctors.

The Nature of Neuroscience Applications

Contrasting work on artificial retinas with that on the science of neurodegeneration and addiction illustrates the two primary aspects of neuroscience applications: (1) a scientific aspect that focuses on identifying relevant brain circuits and understanding how these function and compute; and (2) an engineering aspect that is focused on how to safely use devices to stimulate the relevant brain circuits to create the desired responses.

FIGURE 6.3 Opioids prescribed by physicians

Source: iStock.com / Johnrob

As previously noted, there is much about the brain's anatomy, physiology, and chemistry that is still not well understood, and addressing the theoretical issues in neuroscience is almost exclusively the purview of academia rather than of industry. There are industrial research programs that tackle basic biological questions in neuroscience, but these are tied to solving problems with a profit motive—usually the development of new drugs.

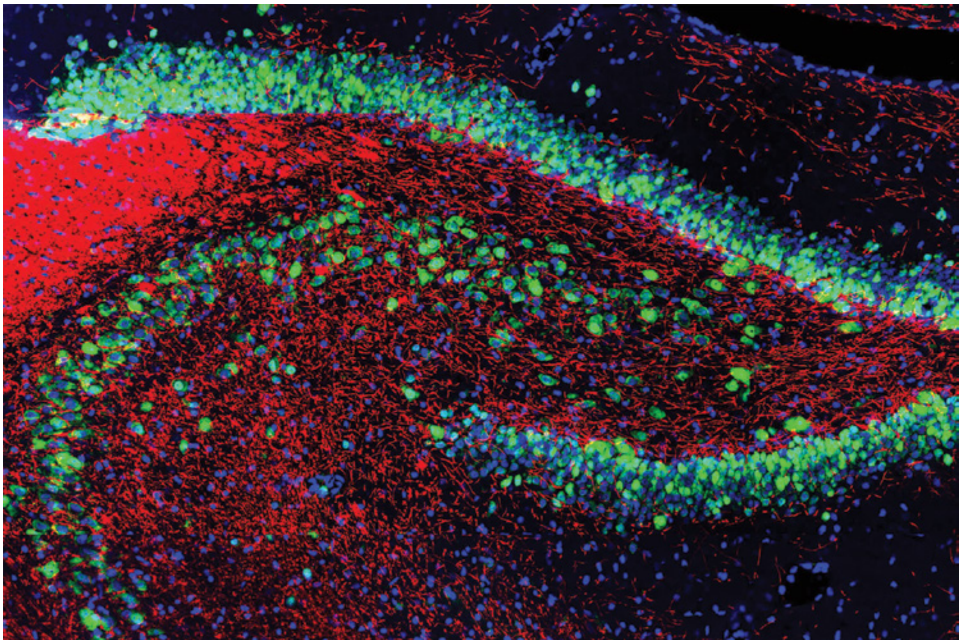
Once the basic science has been developed and a research area approaches an economically viable application, industry does a much better job of developing it. Consequently, helping to smooth the friction of moving a project from academia to industry is crucial to overcoming roadblocks in development. Incubators and accelerators can help transition the findings of basic research to applications by aiding in high-throughput screening—the use of automated equipment to rapidly test samples—and in prototyping. With viable prototypes, new companies can be created or licenses granted to existing companies to produce a final product. Such activities are critical in facilitating the integration of well-understood scientific theory and technical engineering into final applications.

Over the Horizon

Molecular and Genetic Atlases of the Brain

Recent technological advances are transforming neuroscience, enabling scientists to create "bottom-up" maps of the brain at the molecular and genetic level rather than starting from high-level neuroanatomy and behavior (see figure 6.4). These new brain atlases integrate detailed maps of neuronal wiring that show where different genes are activated in the brain and also highlight electrical recordings from different brain regions, allowing scientists to compare differences between healthy and diseased

FIGURE 6.4 A rat hippocampus



Source: Gerry Shaw, Wikimedia Commons, 2015, CC BY-SA 3.0

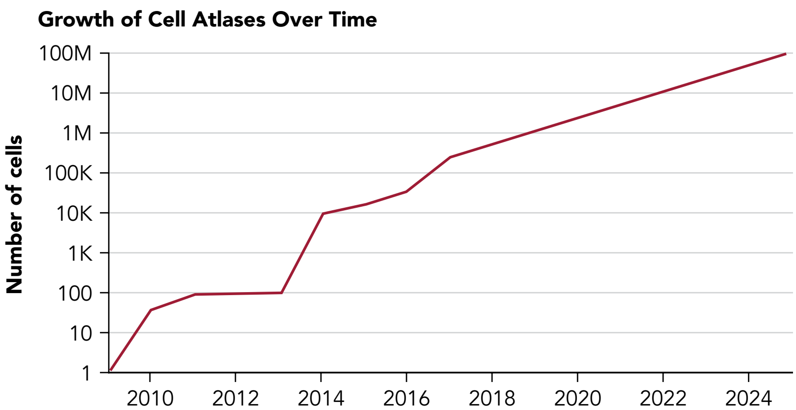
brains. These atlases have already revealed key insights into the genetic mechanisms of Alzheimer's and Parkinson's disease.[15] Understanding the genetic mechanisms behind these diseases could enable the development of rationally designed therapeutics to strengthen brain resilience. (Rational drug design refers to understanding the specific biological target of a disease, such as a protein, and then designing a molecule that precisely fits and interacts with it.)

Over the past decade and a half, we have advanced from mapping individual cells to creating atlases of 100 million cells in a single scientific study,[16] thus accelerating scientific discovery (see figure 6.5).

Beyond facilitating insights into Alzheimer's and Parkinson's, brain atlases have the potential to become foundational "molecular observatories" for neuroscience, providing a shared resource for academic and industry researchers to chart new hypotheses and therapeutic targets for a broad range of neurological diseases. Coupled with substantial decreases in the cost of DNA sequencing (in some cases by several orders of magnitude), gene synthesis, and computing capabilities, these new molecular and genetic approaches to studying the brain have been enabled by the National Institutes of Health (NIH) Brain Research Through Advancing Innovative Neurotechnologies (BRAIN) Initiative.

Currently, it costs about $2 million to $4 million to create atlases with 100 million cells.[17] However, if the aforementioned technological trends continue with federal funding, routine atlassing of the human and the mouse brain at a cost below $10,000 could be possible in the next ten years. Continued investment in these foundational technologies and datasets has the promise to accelerate both fundamental research and translational drug discovery (i.e., the process of moving biomedical knowledge from

FIGURE 6.5 There has been an exponential increase in the size of cellular atlases over the past sixteen years



Source: See Valentine Svensson, Roser Vento-Tormo, and Sarah A Teichmann, "Exponential Scaling of Single-Cell RNA-Seq in the Past Decade," *Nature Protocols* 13, no. 4 (2018): 599–604, https://doi.org/10.1038/nprot.2017.149; and Jesse Zhang et al.,"Tahoe-100M: A Giga-Scale Single-Cell Perturbation Atlas for Context-Dependent Gene Function and Cellular Modeling" bioRxiv 2025.02.20.639398, doi: https://doi.org/10.1101/2025.02.20.639398 the laboratory to patient care and public health for developing new, effective, and accessible therapies for diseases).

Organoid Models of Human Brain Development and Disease

Developing new treatments for neurological diseases requires an understanding of underlying disease mechanisms and an ethical way to test potential therapies. Since testing new therapeutics in humans poses ethical challenges, scientists must first use model organisms, such as mice, to conduct experiments. However, several neurodegenerative diseases, such as Alzheimer's, are specific to humans and poorly replicated in other species. Organoid models serve as a promising alternative, enabling human-specific disease research without direct experimentation on patients.

Organoid models, or organoids, are three-dimensional cellular structures that self-assemble under specific culture conditions to replicate key aspects of human tissue.[18] Stem cells are derived from individual patients: Their cells—often skin cells—are harvested and then reprogrammed into stem cells, which are then differentiated into the desired cell type (e.g., neurons) that makes up an organoid. These organoids can then be studied in a dish to better understand the biology of a disease or to screen potential therapies that can slow a disease's progression. One of the most promising features of organoid models is that they can be personalized to individual patients. This offers unique advantages in several areas, such as, for example, better understanding and treatment of rare diseases.

Scientists can also transplant organoids into xenografted mice—genetically immunocompromised rodents with integrated human cells—to study disease processes in vivo. This is often done to explore how environmental conditions affect neurodegenerative diseases (e.g., how physical activity influences the progression of amyotrophic lateral sclerosis, also known as Lou Gehrig's disease).[19] Such exploration cannot be done in a dish.

Recent guidance from the US Food and Drug Administration has recommended reducing the use of mice in drug testing,[20] opening the door for new approaches, such as organoid models and xenografted mice, to prove therapeutic safety and efficacy. However, there are not yet clear legal and ethical standards for research involving xenografted mice. This is an area where collective discussions between scientists can help to provide greater clarity. While using organoids is becoming a very useful way to study human cells in a dish, it remains an artificial system and is unlikely to replace the role of mice in the understanding of neuroscience and beyond.

Alzheimer's Disease Detection and Treatment

The potential for early detection prior to the onset of cognitive impairment is higher than it has ever been before. Current-generation diagnostic tools now include the ability to cheaply test for biomarkers from blood plasma paired with more accurate but expensive spinal taps and positron emission tomography (PET), which scans for toxic tau and amyloid buildup. A rollout of mass blood-plasma screening, along with confirmation using more expensive tests, might mean the anti-amyloid drugs could be applied before cell death and clinical symptoms manifest themselves, possibly increasing their effectiveness.

Neuroscience and AI

As understanding of the mathematics of our neural computations increases, these computational models may have direct relevance to AI. In particular, machine learning requires vast training datasets. By contrast, humans can learn languages with a small fraction of the training data that AI models require (for more discussion of this point, please refer to chapter 1, on artificial intelligence). A better understanding of the mathematical principles that define how human brains compute may therefore improve AI. The melding of neuroscience theory and AI is a topic of increasing interest under the umbrella of Stanford's Wu Tsai Neurosciences Institute.[21]

Policy Issues

Disconnect Between Public Interest and Capability

The brain is perhaps the least understood, yet most important, organ in the human body. Demand for neuroscience research advances and applications—including understanding brain circuitry, developing new drugs, treating diseases and disorders, and creating brain–machine interfaces—is therefore expected to continue to grow considerably over the coming years. The Society for Neuroscience's annual meeting draws close to thirty thousand attendees.[22]

Science fiction and fantastical headlines fuel beliefs that mind-reading technology, brains controlled by computers, and other dystopias are imminent. In reality, work to comprehend the brain's staggering complexity remains in its early stages. Most advances involve incremental progress, expanding our theoretical foundations rather than producing revolutionary leaps to futuristic applications. This vast gap between public expectations and scientific reality creates an environment ripe for exploitation. Impatience for solutions to pressing medical problems like dementia and mental illness leaves many open to dubious proclamations or pseudoscience.

The Impact of Cognitive and Behavioral Neuroscience on Law

Cognitive and behavioral neuroscience, which studies the biological basis of thoughts and actions, has broad implications for public policy. For example, a basic aspect of criminal law is the nature and extent of an individual's responsibility for a criminal act. Under a 2005 US Supreme Court ruling, minors under eighteen years of age cannot be subject to the death penalty for crimes they have committed, because adolescent brains are not fully developed, which puts minors at higher risk of impulsive, irrational thoughts and behaviors.[23]

Funding Cuts to Transformative Neuroscience

Over the past decade, much of the work outlined in this chapter was funded by the BRAIN Initiative.

Science fiction and fantastical headlines fuel beliefs that mind-reading technology, brains controlled by computers, and other dystopias are imminent. In reality, work to comprehend the brain's staggering complexity remains in its early stages.

Starting in 2014, this initiative aims to be the equivalent of the Human Genome Project for the human brain. Research from the BRAIN Initiative has helped neuroscience generate advances that aid specifically in translating science to medicine. In 2024, however, the initiative's budget was cut by 40 percent, from $680 million to $402 million. The decline was due to a combination of reduced funding from the NIH and through the 21st Century Cures Act. Funding through that legislation fell by an additional $81 million in 2025.[24] Without additional financial support through the NIH, neuroscience research in the United States and the country's ability to tackle some of the most societally impactful diseases will decline.

Foreign Collaboration

Human expertise will continue to be the primary driver of future advances in neuroscience, and success will continue to depend on the United States being the best place for international scientists to train, conduct research, and use their own expertise to teach the next generation of scientists. Against this backdrop, the apparent targeting of US scientists with personal and professional links to China raises concerns,[25] and the United States only loses if these scientists leave and move their labs to China.

Another concern is intellectual property protection. Nearly 80 percent of US biotech companies currently outsource research to Chinese biotechnology firms, such as WuXi AppTec, which offer lower labor costs and benefit from an expanding domestic scientific workforce.[26] Future US leadership in biotechnology will require carefully managing intellectual property risk;[27] increasing investment in technologies that improve scientific productivity in the United States; funding fundamental neuroscience research; and retaining the top US-trained, foreign-born scientists as part of the US scientific workforce.

Ethical Frameworks

Neuroscience research naturally raises many ethical concerns that merit careful, ongoing discussion and monitoring. Chief among these is research on human subjects, which is governed by several existing frameworks and regulations that guide neuroscience studies in American academia today. Ethical guidelines for scientific research are usually national, not international. Some countries might allow particular types of brain research and drugs, while others might not; for example, a nation might permit experimentation on prisoners or on ethnic minorities. Managing differences in state research regimes will be critical to harnessing the power of international collaboration.

NOTES

1. Marcus E. Raichle and Debra A. Gusnard, "Appraising the Brain's Energy Budget," *Proceedings of the National Academy of Sciences* 99, no. 16 (July 2002): 10237–39.

2. Frederico A. C. Azevedo, Ludmila R. B. Carvalho, Lea T. Grinberg, et al., "Equal Numbers of Neuronal and Nonneuronal Cells Make the Human Brain an Isometrically Scaled-up Primate Brain," *The Journal of Comparative Neurology* 513, no. 5 (2009): 532–41, https://doi.org/10.1002/cne.21974.

3. Stanford Medicine, "The Stanford Artificial Retina Project," accessed August 30, 2023, https://med.stanford.edu/artificial-retina.html.

4. Winston Wong, "Economic Burden of Alzheimer Disease and Managed Care Considerations," *American Journal of Managed Care* 26, no. 8 (2020): S177–83, https://doi.org/10.37765/ajmc.2020.88482.

5. Alzheimer's Association, "2023 Alzheimer's Disease Facts and Figures," March 14, 2023, https://doi.org/10.1002/alz.13016; A. W. Willis, E. Roberts, J. C. Beck, et al., "Incidence of Parkinson's Disease in North America," *Parkinson's Disease* 8, no. 170 (2022), https://doi.org/10.1038/s41531-022-00410-y.

6. Vassilis E. Koliatsos and Vani Rao, "The Behavioral Neuroscience of Traumatic Brain Injury," *Psychiatric Clinics of North America* 43, no. 2 (2020): 305–30, https://doi.org/10.1016/j.psc.2020.02.009.

7. Wu Tsai Neurosciences Institute, "NeuroChoice Initiative (Phase 2)," Stanford University, accessed August 30, 2023, https://neuroscience.stanford.edu/research/funded-research/neurochoice.

8. Low end: "The High Price of the Opioid Crisis, 2021," infographic, the Pew Charitable Trusts, August 27, 2021, https://www.pewtrusts.org/en/research-and-analysis/data-visualizations/2021/the-high-price-of-the-opioid-crisis-2021. High end: Feijin Luo, Mengyao Li, and Curtis Florence, "State-Level Economic Costs of Opioid Use Disorder and Fatal Opioid Overdose—United States, 2017," *Morbidity and Mortality Weekly Report* 70, no. 15 (2021): 541–46, http://dx.doi.org/10.15585/mmwr.mm7015a1.

9. "National Drug Threat Assessment, 2011," Office of Justice Programs, US Department of Justice, August 2011, https://www.ojp.gov/ncjrs/virtual-library/abstracts/national-drug-threat-assessment-2011.

10. "Drug Overdose Deaths: Facts and Figures," National Institute on Drug Abuse, National Institutes of Health, August 2024, https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates.

11. Overdose included in the accidental death statistic in "Leading Causes of Death," National Health Statistics, US Centers for Disease Control and Prevention, last modified May 2, 2024, https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm.

12. "Provisional Drug Overdose Death Counts," Vital Statistics Rapid Release, National Center for Health Statistics, US Centers for Disease Control and Prevention, accessed August 26, 2025, https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm.

13. "FDA Takes Steps Aimed at Fostering Development of Non-Addictive Alternatives to Opioids for Acute Pain Management," US Food and Drug Administration, February 9, 2022, https://www.fda.gov/news-events/press-announcements/fda-takes-steps-aimed-fostering-development-non-addictive-alternatives-opioids-acute-pain-management.

14. Gordy Slack, "Social Aversion During Opioid Withdrawal Reflects Blocked Serotonin Cues, Mouse Study Finds," Wu Tsai Neurosciences Institute, Stanford University, November 2, 2022, https://neuroscience.stanford.edu/news/social-aversion-during-opioid-withdrawal-reflects-blocked-serotonin-cues-mouse-study-finds.

15. Mariano I. Gabitto, Kyle J. Travaglini, Victoria M. Rachleff, et al., "Integrated Multimodal Cell Atlas of Alzheimer's Disease," *Nature Neuroscience* 27, no. 12 (2024): 2366–83, https://doi.org/101038/s41593-024-01774-5; Tushar Kamath, Abdulraouf Abdulraouf, S. J. Burris, et al., "Single-Cell Genomic Profiling of Human Dopamine Neurons Identifies a Population That Selectively Degenerates in Parkinson's Disease," *Nature Neuroscience* 25, no. 5 (2022): 588–95, https://doi.org/10.1038/s41593-022-01061-1.

16. Fuchou Tang, Catalin Barbacioru, Ying Wang, et al., "mRNA-Seq Whole-Transcriptome Analysis of a Single Cell," *Nature Methods* 6 (2009): 377–82, https://doi.org/10.1038/nmeth.1315; Jesse Zhang et al., "*Tahoe-100M*: A Giga-Scale Single-Cell Perturbation Atlas for Context-Dependent Gene Function and Cellular Modeling," preprint, bioRxiv, February 20, 2025. https://doi.org/10.1101/2025.02.20.639398.

17. Anna Elz et al., "Evaluating the Practical Aspects and Performance of Commercial Single-Cell RNA Sequencing Technologies," preprint, bioRxiv, May 19, 2025, https://doi.org/10.1101/2025.05.19.654974.

18. Zixuan Zhao, Xinyi Chen, Anna M. Dowbaj, et al., "Organoids," *Nature Reviews Methods Primers* 2, no. 1 (2022): 94, https://doi.org/10.1038/s43586-022-00174-y.

19. Laura Chapman, Johnathan Cooper-Knock, and Pamela J. Shaw, "Physical Activity as an Exogenous Risk Factor for Amyotrophic Lateral Sclerosis: A Review of the Evidence," *Brain* 146, no. 5 (2023): 1532–47, https://doi.org/10.1093/brain/awac470.

20. "FDA Announces Plan to Phase Out Animal Testing Requirement for Monoclonal Antibodies and Other Drugs," press release, US Food and Drug Administration, April 10, 2025, https://www.fda.gov/news-events/press-announcements/fda-announces-plan-phase-out-animal-testing-requirement-monoclonal-antibodies-and-other-drugs.

21. "Center for Mind, Brain, Computation, and Technology," Wu Tsai Neurosciences Institute, Stanford University, accessed October 14, 2024, https://neuroscience.stanford.edu/initiatives-centers/center-mind-brain-computation-and-technology.

22. "Attendance Statistics: Meeting Attendance," Society for Neuroscience, accessed August 30, 2023, https://www.sfn.org/meetings/attendance-statistics.

23. Roper v. Simmons, 543 U.S. 551 (2005), https://supreme.justia.com/cases/federal/us/543/551/.

24. "Understanding the BRAIN Initiative Budget," National Institutes of Health, the BRAIN Initiative, last reviewed March 25, 2025, https://braininitiative.nih.gov/funding/understanding-brain-initiative-budget.

25. Jeffrey Mervis, "Pall of Suspicion: The National Institutes of Health's 'China Initiative' Has Upended Hundreds of Lives and Destroyed Scores of Academic Careers," *Science*, March 23, 2023, https://www.science.org/content/article/pall-suspicion-nihs-secretive-china-initiative-destroyed-scores-academic-careers.

26. Jared S. Hopkins and Clarence Leong, "U.S. Drugmakers Are Breaking Up with Their Chinese Supply-Chain Partners," *The Wall Street Journal*, November 1, 2024, https://www.wsj.com/health/pharma/china-manufacturing-astrazeneca-supply-chain-2ddecb11.

27. Michael Martina, Michael Erman, and Karen Freifeld, "Exclusive: China's WuXi AppTec Shared U.S. Client's Data with Beijing, U.S. Intelligence Officials Told Senators," *Reuters*, March 28, 2024, https://www.reuters.com/technology/chinas-wuxi-apptec-shared-us-clients-data-with-beijing-us-intelligence-officials-2024-03-28/.

STANFORD EXPERT CONTRIBUTORS

Dr. Kang Shen
SETR Faculty Council, Frank Lee and Carol Hall Professor of Biology, and Professor of Pathology

Dr. Keith Humphreys
Esther Ting Memorial Professor in the Department of Psychiatry and Behavioral Sciences

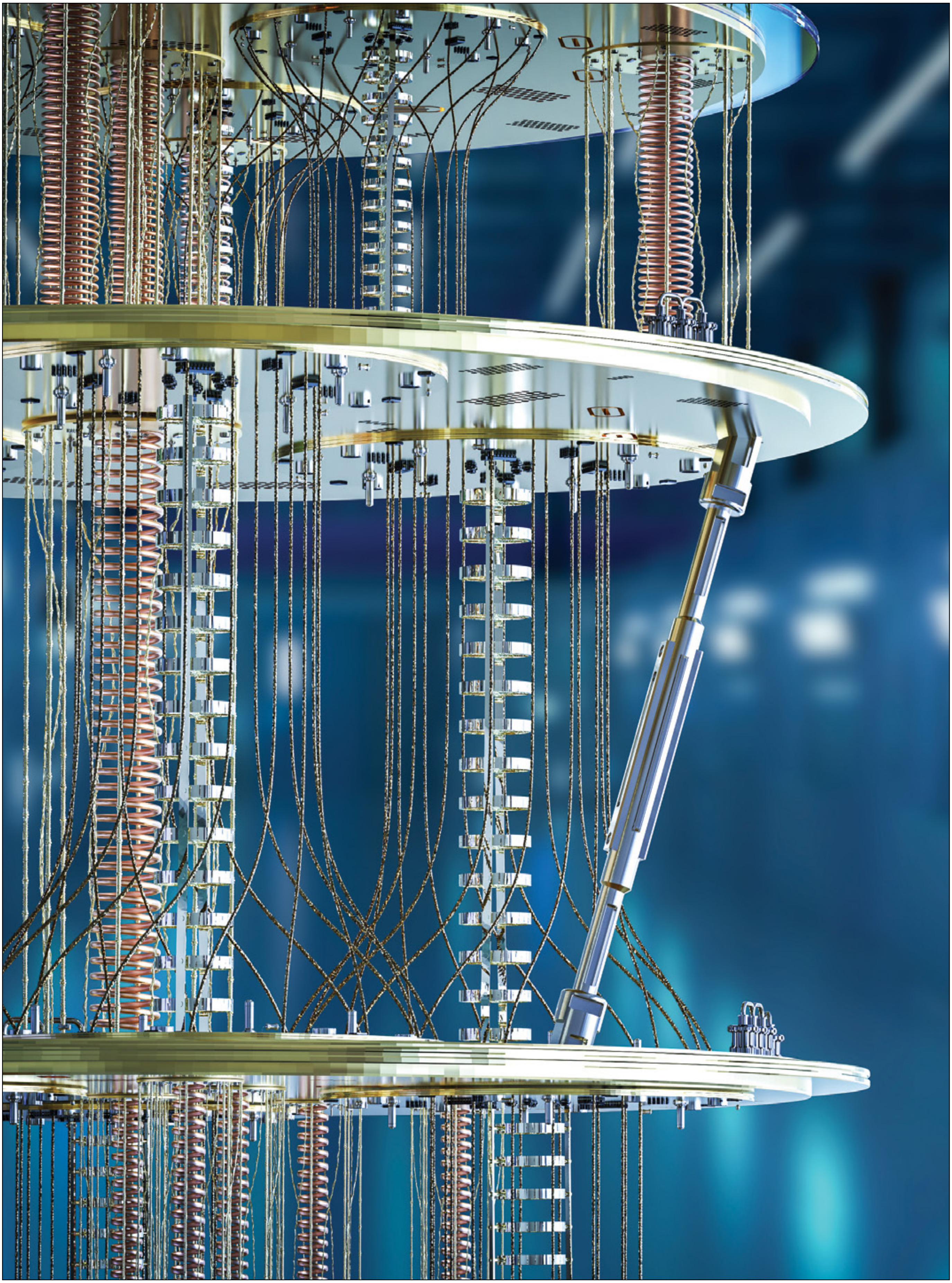
Dr. Paul Nuyujukian
Assistant Professor of Bioengineering, of Neurosurgery, and, by courtesy, of Electrical Engineering

Dr. Michael Greicius
Iqbal Farrukh and Asad Jamal Professor and Professor, by courtesy, of Psychiatry and Behavioral Sciences

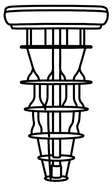
Dr. Will Allen
Assistant Professor of Developmental Biology

Dr. Sergiu Pasca
Kenneth T. Norris, Jr. Professor of Psychiatry and Behavioral Sciences

Thomas Lau
SETR Fellow and PhD Student in Bioengineering



07



QUANTUM TECHNOLOGIES

KEY TAKEAWAYS

- Quantum computing is advancing rapidly, making clear progress toward solving practical problems such as breaking existing public-key encryption algorithms, enabling new materials design, and supporting applications in chemistry. More speculative uses include machine learning, weather modeling, and financial portfolio optimization.

- Quantum networking and sensing are emerging as powerful technologies—networking may be critical for scaling computers to utility levels, while sensors are already transforming fields such as medical imaging and gravitational detection.

- Government-funded basic research in academic labs remains the foundation for breakthroughs, and sustained investment is essential to maintain leadership as companies push applications toward real-world utility.

Overview

Quantum technologies are based on the physics of quantum mechanics, which emerged early in the twentieth century. Since then, quantum mechanics has shaped many technologies, from nuclear weapons to the transistors in smartphones to magnetic resonance imaging (MRI) machines in hospitals. But in these applications, the constituent atoms have been controlled in aggregate, with large, uncontrolled groups of particles all in multiple states manipulated together as an ensemble.

Modern quantum technology seeks to control the components of an ensemble particle by particle. In 2025, the Nobel Prize in Physics was awarded for "the discovery of macroscopic quantum mechanical tunnelling and energy quantization in an electric circuit."[1] The goal is to develop real-world applications by precisely controlling many particles that are all doing many things simultaneously, which requires an enormously complex effort. In the past twenty years, several different hardware approaches have emerged to control and detect individual particles and their states. In just the past five years, this control has become strong enough that useful quantum technologies are starting to be built and used (see sidebar on entanglement).

Precise control allows for the management of superposition and entanglement. Superposition is an empirically validated principle of quantum mechanics stating that particles can be in different states (e.g., in different places or spinning different ways) simultaneously. Entanglement is the quantum phenomenon in which particles are linked in a correlational sense. This means that measuring a property of one reveals a correlated property of the other instantly, even if they are separated by very long distances and their properties have not been individually determined beforehand (see sidebar).

A particularly important aspect of this new quantum paradigm is that measuring a quantum system disturbs it. Therefore, it does not work to measure everything and process the data later. Instead, it is necessary to carefully craft new types of hardware that measure directly and exclusively what we want to know and do so only when we want to know it. Violating any of these principles (e.g., by making a query at the wrong time) reduces the quantum-enhanced accuracy in measuring items of interest. This insight is important to all of the technologies discussed below.

While there are many potential technologies based on quantum principles, the three most mature are **quantum computing**, **quantum communication**, and **quantum sensing**. In the long run, it is most likely that these quantum technologies will complement rather than replace their classical counterparts.

- **Quantum computing** will be useful primarily for solving problems that classical computing cannot, but these problems for the most part will be niche problems rather than general ones of broad interest.

ON ENTANGLEMENT

Alice and Bob share a coin cut perfectly into two halves, one showing heads and the other tails. Alice keeps the tails half and places the heads half into an opaque envelope, which she gives to Bob without revealing its content. From Bob's perspective, before opening the envelope, the coin inside could be either heads or tails, resembling a quantum superposition. When he eventually observes heads, he instantly knows Alice's half must be tails, no matter the distance between them.

This correlation illustrates the essence of quantum entanglement: Bob's measurement does not cause Alice's outcome but instead reveals an existing relationship that manifests instantly without any transfer of information faster than light.

Although the coin analogy is helpful for understanding entanglement, it is not perfect. In the analogy, the coin halves have definite states regardless of whether they are observed—the outcome is simply hidden until revealed. In contrast, entangled quantum particles do not have definite states before measurement. Instead, their joint state exists as a superposition, meaning that the properties of both particles are fundamentally linked and undefined until one is measured. When a measurement is made on one particle, the outcome of a corresponding measurement on its partner is instantly determined—no matter how far apart they are.

This behavior cannot be fully understood using everyday, classical intuitions. It is a uniquely quantum phenomenon, confirmed by decades of rigorous experiments. Importantly, this "spooky" correlation is not just a scientific curiosity or a matter of philosophy—it is the foundation for groundbreaking quantum technologies in computing, communication, and sensing.

[Quantum] behavior cannot be fully understood using everyday, classical intuitions.

- **Quantum communication** may also have niche applications, such as for cryptographic key distribution and distributed quantum computing; these are discussed later in this chapter. However, it is unlikely to be a broadly applicable technology for communications infrastructure because it requires specialized hardware to implement.

- **Quantum sensors** will not render classical sensors obsolete in the short to medium term. Rather, they will be used primarily in application areas where they have particular advantages, such as greater sensitivity or measurement stability. To the extent that quantum sensors push the limits of what quantum mechanics allows in terms of power usage, sensitivity, size, and so on, most sensors are likely to eventually incorporate quantum technology in some form (though not necessarily as quantum networked sensors).

Quantum Computing

Essential Points

- Quantum computers promise major speedups for factoring large numbers and simulating quantum particles and processes—essential for breaking certain kinds of encryption and for chemistry and materials science. Quantum computing will remain a specialized, niche technology rather than one with broad real-world impact unless and until practical and effective algorithms for finance and other fields are developed.

- All known applications of quantum computers will require error correction to outperform traditional, classical computers. Quantum computing hardware has reached the "break-even" point (i.e., where error correction is feasible) enabling practical scaling.

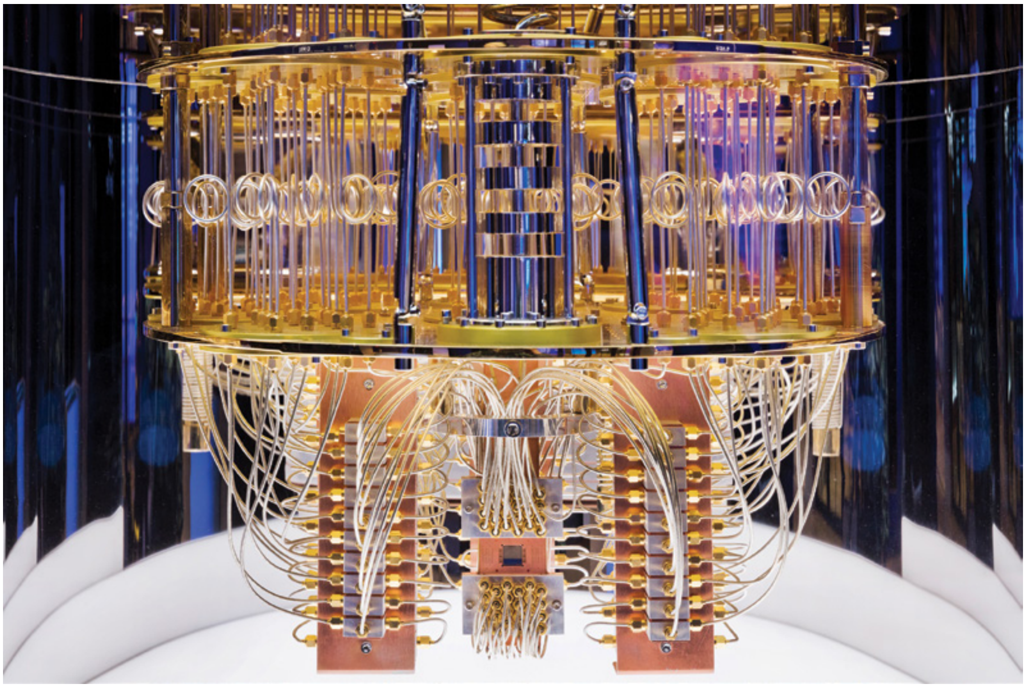
- The main scaling challenge is not simply in increasing the number of quantum bits, or qubits, in a module. Rather it is in controlling individual qubits. Scaling of quantum computers comparable to what has been achieved for classical computers depends on the availability of robust error-correction approaches—expected within the next few years—and on leveraging established techniques from semiconductor and photonics engineering.

About Quantum Computing

Quantum technology offers a fundamentally new computational paradigm. Classical computers use individual bits as the smallest unit of information, with each being 0 or 1. In contrast, qubits—the smallest unit of information for a quantum computer—can be in multiple states simultaneously; that is, qubits can exist in superposition. This reality is one aspect of what allows quantum computers (like the one shown in figure 7.1) to process a vast number of possibilities at once, a phenomenon called quantum parallelism. This capability makes quantum machines a potentially game-changing advance in the field of computing.

In gate-based quantum computing—in contrast to quantum simulation, discussed later in this chapter—qubits are manipulated through discrete operations performed by quantum gates. These gates alter the state of the qubits they act upon. Serving as the quantum analog of logic gates in classical computers, they form the fundamental building blocks of quantum circuits that execute calculations by chaining together individual gate operations.

FIGURE 7.1 The wiring infrastructure of a quantum computer

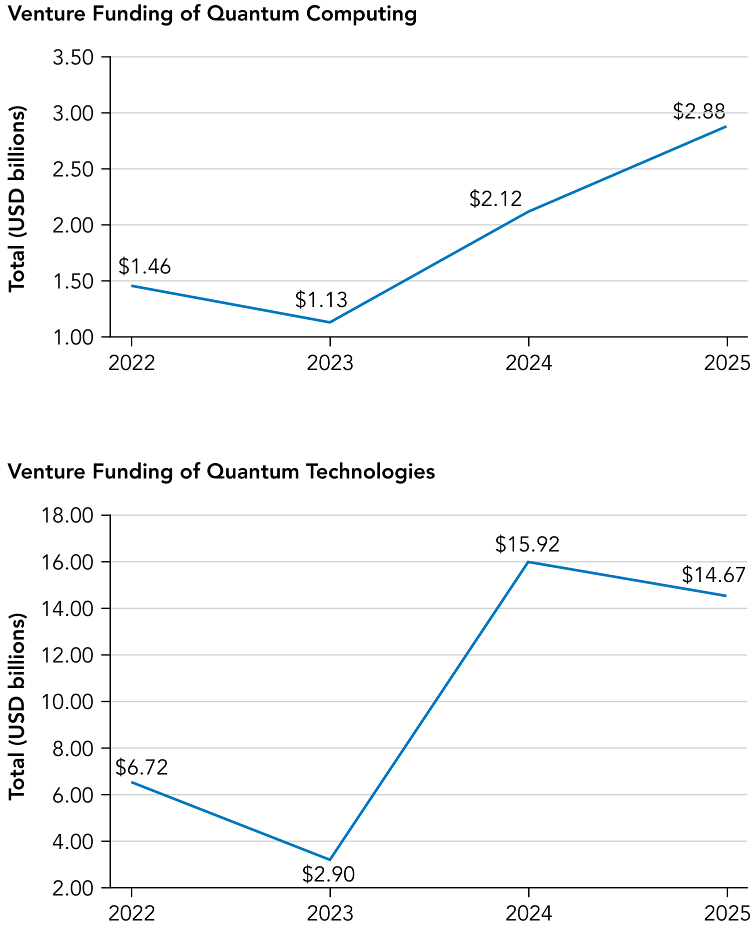


By carefully sequencing these operations, gate-based quantum computers can, in principle, perform a very specific set of calculations, including for quantum chemistry and certain kinds of codebreaking, much faster than their classical counterparts. This possibility and the potential for a broader array of applications down the line have driven up public and private investment in quantum computing in recent years. (Figure 7.2 shows statistics on venture capital investment.) This rapid growth reflects investor interest in the field, though significant challenges must be resolved before quantum computing's value can be realized.

Chief among these challenges is obtaining the answer to a problem once a quantum computation has been performed. Because the qubits are in a superposition state, a computation that is performed on them generates many possible results. However, nearly all of these will not be useful to solving the problem at hand. Realizing the advantage of a quantum computer requires developing algorithms that surface the useful result often enough that it's possible to identify it without having to repeat the calculation too many times. This is fundamentally different from classical computing, which produces deterministic results of computations.

A second important challenge is that quantum computing operations can be disrupted by errors caused by small amounts of "noise" in the environment, such as atomic decay, small vibrations, or tiny changes in temperature. This noise can disrupt the delicate quantum state of qubits and lead to errors. In other words, interaction with the outside world generally destroys superposition states that enable quantum parallelism, so it's crucial to minimize this noise as much as possible and to correct the errors that cannot be avoided.

FIGURE 7.2 Funding trends in quantum computing and quantum technologies



Note: 2025 numbers are partial-year data.

Source: CB Insights

Compensating for these errors—a process called error correction—accounts for upward of 90 percent of what a quantum computer actually does; the remaining 10 percent, or less, is spent actually computing. Error correction is the process of storing information in a way that it is resilient to noise and other error mechanisms. This encoding uses many physical qubits in the quantum hardware to encode one logical qubit, which is a qubit that is robust against noise. A quantum computer performs its algorithms between these logical qubits while continuously correcting the errors within them that emerge through the processing. With current approaches, hundreds to thousands of physical qubits are needed to encode every logical qubit used for computation.

A third challenge is the fact that today, there are few problem-solving algorithms to run on quantum computers that have any practical value. The quantum research community still lacks a complete understanding of the algorithms that would ensure quantum computing provides any speedup at all over classical computers.

Put differently, an important constraint on accomplishing useful tasks with quantum computers is the current lack of suitable algorithms for those purposes; this is in addition to the lack of sufficiently powerful quantum computers themselves. The lack of algorithms is partly due to a chicken-and-egg problem: It is extremely challenging to invent and test new algorithms without the underlying hardware to rapidly test and guide the process. Error-corrected hardware, when available, will enable empirical development and benchmarking of quantum algorithms; currently, this can be done only by formal mathematical proofs of correctness and speed.

Key Developments

QUBIT TECHNOLOGY

Numerous hardware approaches are under active exploration for construction of physical qubits. At the time of this writing, the most visible candidates include the following:

- **Trapped ions** Atoms that are missing an electron (i.e., ions) are not electrically neutral, so they can be trapped by radio-frequency electric fields. The ion encodes information in its internal states, and the information can be read out using laser beams. Computations are performed through conducting controlled collisions between ions.

- **Superconducting circuits** Composed of nanofabricated chips, these circuits must be kept cold at temperatures near absolute zero (–273.15°C). Like traditional electronics, computations are performed by running currents through the chips. However, the fabrication procedures are different from those of ordinary semiconductors.

- **Neutral atoms** These are atoms that are not missing any electrons and hence must be trapped with laser beams. Moving these laser beams around allows the atoms to collide and perform computations. The laser traps are weak, so the atoms must be held in an extreme vacuum that is comparable to interplanetary space.

- **Silicon spin qubits** These are single electrons trapped in structures resembling modern silicon transistors that often use standard semiconductor foundry processes. These qubits also operate at temperatures near absolute zero. They are somewhat less developed than the approaches described above due to their extreme sensitivity to small material defects. However, they are advancing rapidly and have the potential to take advantage of the large investments and rapid learning rates in the semiconductor industry.

- **Photonic qubits** These qubits encode information in states of optical frequency light fields (usually telecom-band photons), perform operations running at room temperature, and interface naturally with fiber networks. Proponents claim scalability through wafer-scale photonics, low-loss fiber, and other factors. In practice, progress has been constrained by the inability to develop an appropriate test environment that could provide convincing evidence of practical advantage.

In the early days of quantum computing, trapped ions and superconducting circuits were seen as the leading candidates. In the past decade, neutral atom quantum technology has become a true competitor to them. This development demonstrates that the "horse race" for the quantum transistor remains open, leaving opportunities for the emergence of alternative approaches that scale more easily and deliver improved underlying performance.

Two dark horses in this category are photonic and topological qubits. Photonic qubits aim to leverage silicon photonics manufacturing to achieve fault tolerance, but using them in quantum gates is technically challenging. Topological qubits aim to store information in a naturally protected form that, under ideal conditions, should be less sensitive to changes in the external environment—though recent research on this technology has been challenged.[2] Both platforms have attracted significant public and private investment and have promised compelling pathways to leapfrog current leaders in demonstrated quantum capabilities. However, neither platform has yet demonstrated computation at even the scale of a few qubits, and both face substantial technical roadblocks.

One of the primary determinants of front-runner status of any physical qubit technology is its fidelity. Fidelity refers to how accurately a quantum gate executes its intended operation compared to an ideal error-free gate—the greater the similarity between the two, the higher the fidelity of the gate under examination. This is important because the computation needed to perform error correction in a quantum computer is itself vulnerable to error.

For a given qubit technology and error-correcting algorithm, a process called "co-design" for the code and platform establishes a break-even threshold for fidelity. This means that the errors corrected by the error-correction circuitry are barely equal in number to the errors introduced because of the operation of that circuitry. Higher gate fidelities increase the former number relative to the latter (more and more errors corrected compared to the number of errors introduced). In practice, this break-even threshold across a number of technologies is around 99 percent.[3]

From a practical standpoint, there is a broad consensus in the field that gate fidelities above 99.9 percent (that is, significantly better than the break-even threshold) will enable the construction of practical quantum computers that do not require a prohibitive amount of error correction or a prohibitive number of physical qubits to implement that error correction.

The most important development in qubit technology is that trapped ions, neutral atoms, and superconducting circuits have all now crossed the "break-even" threshold. However, quantum computing faces uncertainty over which of these—or perhaps one of the dark horse possibilities—will lead to the best scalable architecture. Unlike classical computing, in which transistors proved superior to vacuum tubes and became the technological foundation for a fabrication industry that set classical semiconductors on the path of rapid and sustained cost reductions over decades,[4] the "quantum transistor" has yet to be identified. Indeed, rather than a single winner, multiple platforms may well coexist, each with unique strengths and challenges. As a result, quantum computing progress is likely to depend on hardware innovation, error correction, and application-specific advances.

QUANTUM MEMORY

Memory in quantum computers serves approximately the same function as it does in classical computers: It holds quantum information in qubits while preserving their quantum properties like superposition and entanglement until operations can be performed on that information. An example of recent

Unlike classical computing, in which transistors proved superior to vacuum tubes . . . the "quantum transistor" has yet to be identified. Indeed, rather than a single winner, multiple platforms may well coexist.

progress is the development of quantum memory that extends storage times of quantum states by a factor of up to thirty.[5] Quantum memory also plays a critical role in quantum repeaters, which are discussed in the Quantum Communications section later in the chapter.

SCALE-UP OF QUANTUM COMPUTERS

At present, the state of the art is computers with thousands of atomic qubits and hundreds of superconducting and ionic qubits. In each of these platforms, these physical qubits have been combined to redundantly encode dozens (currently up to forty-eight) logical qubits,[6] albeit at error rates orders of magnitude too high for scalable computation. To achieve the error rates and computer sizes necessary for scalable computing, more logical qubits to perform computations and more encoding redundancy to reduce the error rates of the logical qubits are required. In other words, many, many more physical qubits are going to be needed.

Current resource estimates indicate that thousands of logical qubits, and thus millions of physical qubits,[7] will be required to build quantum computers that can break current public-key encryption schemes. Achieving systems at this scale requires innovation in qubit design and performance, and substantial progress in the ability to manipulate so many qubits simultaneously.

Thus, from here, the game is to scale, increasing the qubit count without sacrificing qubit performance. As mentioned above, a gate fidelity of 99.9 percent (corresponding to an error rate of 0.1 percent per gate operation) is expected to be sufficient for quantum computation at scale. Along similar lines, a million physical qubits, either distributed across networked modules or in a single large quantum computer, will be required to perform practically useful computation.[8] Present-day industry road maps suggest that individual quantum computing modules with tens of thousands to hundreds of thousands of physical qubits are feasible within the next decade or so.[9]

Controlling the operations in which these qubits are involved is another key task in scaling up. A quantum computer must be able to control each qubit it uses. To do so, a control channel needs to be established that sends signals to every qubit; if a quantum computer with a million qubits requires each individual qubit to be manipulated on a microsecond timescale, a control bandwidth of around one terabit per second is needed.

Such rates are usually achieved, but they exceed the internal memory bandwidth of a standard microprocessor chip. Both the processing power to generate this control and the physical control hardware require technology development. This will likely be in the form of applications-specific integrated circuits for all quantum computing platforms, faster cameras, and other supporting technologies. Today, control technology for large numbers of qubits lags behind the ability to fabricate and trap them. Developing tools for control is likely to be a broad effort, requiring deep understanding of hardware architectures, significant expertise in various hardware development paradigms, and novel solutions combining them.

QUANTUM SIMULATION

Despite a huge amount of contemporary discussion and excitement about the revolutionary importance of quantum computing, there are currently few real-world applications of the technology. Those that do exist are examples of quantum simulation.

Quantum simulation works by creating a physical system—in this case, physical qubits—whose behavior is mathematically analogous to the problem being solved. The key insight is that the equations governing the physical system must have the same mathematical structure as those governing the problem; the specific physical details are irrelevant. Since the system's evolution is determined by its governing equations (and its starting conditions), the physical realization of the original problem becomes a simulator of it. Observations of the system's behavior can then be interpreted as the problem's solution.

By applying carefully controlled stimuli, such as electromagnetic fields or laser pulses, to qubits and observing their subsequent behavior, it is possible to drive the system to evolve naturally. Quantum simulations are also sensitive to fluctuations in the external environment. However, they typically do not employ active error-correction protocols. Instead, their design emphasizes minimizing external disturbances as much as possible.

Quantum simulation is most suitable for specific, narrowly defined problems. Perhaps the most prominent application is the modeling of the quantum properties of complex materials. This effectively mimics those materials using qubits, thereby allowing researchers to explore material behaviors otherwise exceedingly difficult to study classically.[10]

For example, French research teams have similarly shown how quantum simulation using neutral atoms for qubits can be used to model solvent interactions in proteins, which is critical for advancing drug discovery.[11] Additional work has demonstrated the usefulness of quantum simulations in clarifying the conditions required for high-temperature superconductivity and in refining classical numerical models of material behavior.[12]

Over the Horizon

APPLICATIONS AND SOFTWARE

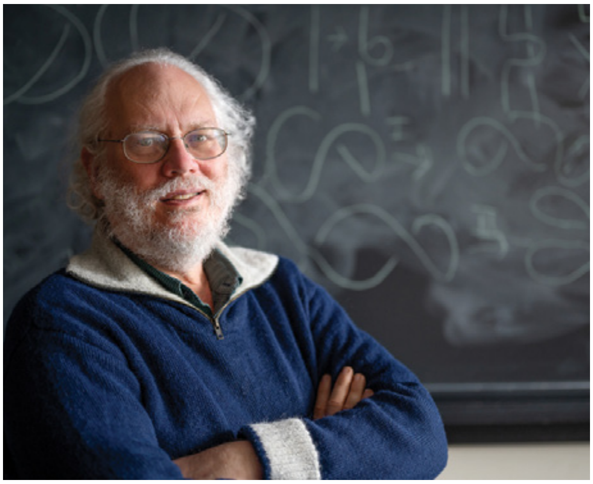
Although gate-based quantum computing could in principle be a general-purpose computing technology, its specialized hardware and software requirements currently make it impractical as one. The expense and effort of meeting such requirements are worth it only when the problems being solved are sufficiently important or have enough economic value.

Below we describe two of the most promising known applications for gate-based quantum computing:

- **Breaking certain algorithms for asymmetric cryptography** Discussed in more detail in chapter 3, on cryptography and computer security, the most commonly used algorithms in asymmetric cryptography are based on the difficulty of factoring large numbers and related problems. Classical computers factor numbers by testing one set of possible factors at a time, a process whose time to completion grows very rapidly as the numbers get very large. Shor's algorithm, developed in 1994 by Peter Shor (see figure 7.3), exploits quantum parallelism to efficiently perform such tests over all possible inputs simultaneously. This reduces the time needed to break encryption from thousands of years to potentially hours or minutes with a sufficiently powerful quantum machine.

  Shor's algorithm is expected to provide exponential speedup over what is possible with classical computing, making it a top candidate application for quantum computing. The best current estimates suggest that the most commonly used asymmetric cryptographic algorithm—RSA-2048—could be breakable in days to weeks on a quantum computer coming online in five to fifteen years.[13] (Note, however, that Shor's algorithm is not an algorithm that works against all possible asymmetric cryptography algorithms. It only works against

FIGURE 7.3 Peter Shor, creator of Shor's algorithm



Source: Christopher Harting, *MIT News*

asymmetric algorithms based on the difficulty of factoring large numbers and related problems.)

- **Simulating the quantum world** From chemistry to materials science, simulations using quantum computers are expected to be exponentially faster than ones using classical computers. They are expected to have significant positive impacts on problems like nitrogen fixation, drug development, and superconductivity[14] and enable breakthroughs in chemical, drug, and catalyst design. The analog simulators currently solving these problems are likely to be replaced by more precise simulations on gate-based quantum computers.[15] Ultimately, quantum computers may generate complex quantum data that classical computers can then use to efficiently manage more routine quantum chemistry simulations.[16]

These application areas have significant economic or national security value, and gate-based quantum computing potentially offers substantial speedups for the above problems over classical computing.

How substantial? Most desirable is an exponential speedup. When this is possible, a quantum computer can finish a task in a practical amount of time that would take a classical computer so long that it's essentially impossible—sometimes longer than the age of the universe. An example would be a classical computer taking a billion years to solve a complex chemistry problem that a quantum computer with exponential speedup could do in a few minutes. When they are possible, exponential speedups are transformational.

By contrast, a polynomial speedup can still be substantial.[17] However, polynomial speedup gains in practice are limited by the much slower clock speeds of quantum computers (that is, the number of operations per second that a quantum computer can perform compared to a classical computer). Equally important, continuing improvements in classical computing hardware over the next decade are likely to deliver performance gains of a similar magnitude and offer a more probable near-term path to achieving them.

For many computational problems that need to be solved, the consensus among experts seems to be that if speedups using quantum computing exist at all, they will most likely be polynomial in nature.[18] In addition, such speedups may require careful optimization of quantum computers and significant hardware development to realize practical gain. Shor's algorithm for factoring (and thus for certain types of codebreaking) and quantum simulation for chemistry and materials science are among the few notable examples of exponential speedups currently known to exist over the best classical algorithm.

In parallel with these algorithmic advances, work is ongoing at a number of companies to develop unified (and often cloud-based) frameworks that are qubit-hardware-agnostic and programmable by non-specialists, thereby making quantum computing more widely accessible. Broad adoption of such tools is predicated on quantum hardware advances that provide fault-tolerant quantum computers and algorithmic advances that clarify commercial use cases.

It is difficult to predict which problems will be most directly impacted by quantum computing. However, what seems clear from the evolution of both classical computers and neural networks is that new computing paradigms always lead to new opportunities. Further research in quantum algorithms is thus essential and is likely to accelerate as quantum computing hardware for testing the algorithms becomes more sophisticated and accessible.

ON QUANTUM SUPREMACY AND ADVANTAGE

While fault-tolerant quantum computers capable of useful computation for codebreaking, quantum chemistry, and other uses remain some years out, an ongoing effort exists to demonstrate *quantum supremacy*. This refers to the quantum computation of any quantity sufficiently complex that a classical computer cannot replicate the same result.

The endeavor to demonstrate quantum supremacy has resulted in an extremely productive race between quantum computing teams and classical computing teams, with the former running ever-more-complicated "random quantum circuits" and the latter demonstrating that they can in fact predict the output of these circuits on classical computers.

The latest generation of superconducting quantum processors are large enough that they can now perform certain calculations that are difficult or impossible to replicate on classical computers.[19] However, this is a benchmark—the calculated quantity is not inherently useful, even if it is beyond the reach of a classical computer to calculate. Nonetheless, this work points to classes of problems where quantum supremacy is indeed possible and drives progress in understanding the quantum/classical computability boundary.

By contrast, the term *quantum advantage* is generally used to denote the superiority of a quantum computer in solving a practical, useful problem faster or more accurately than a classical computer. In addition, it signals a relevance to real-world tasks and potential commercial applicability. To date, true quantum advantage in computing has not been achieved on any useful real-world problem.

Quantum Communication

Essential Points

- Post-quantum encryption algorithms are already being deployed and used to protect against attacks based on factoring and related problems. Quantum communication for key distribution will be broadly useful only if it turns out that these algorithms are flawed in practice.

- Quantum networking for connecting individual quantum computing modules will likely be necessary at least in the shorter term to solve useful or meaningful problems.

About Quantum Communication

Quantum communication uses the principles of quantum mechanics, such as superposition and entanglement, to encode, transmit, and secure information between separate systems. It has two primary applications: One is related to privacy and security in data transmission and identification. The other is transmission of intrinsically quantum data, essential for tasks in scalable quantum computing and networks of sensors linked through quantum entanglement.

Quantum-Enabled Data Security

The security afforded by quantum communication is based on its application to what is known as the key distribution problem, an essential element of secure digital communication.

Today's public-key cryptography, discussed in chapter 3, on cryptography and computer security, is susceptible to attacks from future quantum computers. This is because, as discussed earlier, it relies upon the difficulty of factoring large numbers (or other related problems) for its security. Cryptographers are developing quantum-resistant algorithms—more precisely, algorithms that will resist Shor's algorithm running on quantum computers. However, if they are unsuccessful, alternative key distribution methods, such as quantum key distribution (QKD), will be necessary.

QKD does not rely on the infeasibility of obtaining private keys from public keys. The security it affords is based on the fact that quantum information cannot be copied.[20] That is, it is impossible to create an exact, independent copy of an arbitrary and unknown quantum state—a statement known as the no-cloning theorem. Copying quantum information always perturbs the original in detectable ways. By contrast, classical information can be perfectly copied without perturbation to the original.

If quantum information cannot be copied, it means that it is impossible to eavesdrop on communications conducted with quantum data. Eavesdropping entails a third party—say, someone called Eve—listening in to a communication from Alice to Bob. If that communication is conducted with classical data, Eve can intercept it in transit without Alice's or Bob's knowledge. Interception implies making a copy of the information that is in transit: The original version is what Alice sends to Bob and is in Bob's possession, and the copy is what Eve has after the interception.

But if the communication is conducted with quantum data, the impossibility of copying quantum information means that if Eve attempts to intercept the message, her interaction will perturb the quantum state of it. As a result, Bob will be able to detect Eve's presence either by a failure to receive the information or by observing measurable errors or anomalies in the received data caused by the perturbation. If Bob receives the quantum information without any such disturbance, it indicates that no third party has accessed the message. In this case, Alice and Bob can use the received quantum information to securely establish a shared cryptographic key, which can then be used to protect subsequent communications between them. This is the process known as QKD.

Quantum communication is often regarded as a guarantor of perfect data security. For example, the European Telecommunications Standards Institute asserts that "QKD is secure now and always will be. By enabling provable security based on fundamental laws of quantum physics, QKD remains resilient even to future advances in cryptanalysis or in quantum computing."[21]

This claim is true as far as it goes, but it omits several important points:

- QKD securely distributes only shared keys that can be used with existing symmetric encryption algorithms. The physical endpoints of the communication system must still be secured. QKD does not by itself solve related challenges such as authentication, side-channel attacks, and man-in-the-middle attacks.[22]
- Existing symmetric encryption algorithms using long keys already provide effectively unbreakable protection for data in transit, and Shor's algorithm provides no leverage in breaking symmetric algorithms. Other quantum algorithms provide modest (polynomial) assistance that is still entirely insufficient to break the encryption algorithms in any reasonable time frame.
- As noted earlier, quantum communication in the form of QKD provides a hedge against a failure to develop public-key encryption algorithms resistant to being broken by quantum computers. But by and large, the cryptographic community has considerable confidence that the attempt to develop such algorithms will be successful.
- QKD's security is effective only when two parties share an already-established trust relationship (e.g., through an initial face-to-face meeting in which they share a secret key so they can later verify their identities to each other). Using QKD, strangers without such a relationship will be able to communicate securely only by relying on trustworthy third parties to establish that initial trust relationship.

Quantum Networking

The security benefits for quantum-enabled data security described above will require the ability to transmit and share quantum information at a range of distances: QKD between continents would require establishing links over very long distances. Building multi-node quantum supercomputers, on the other hand, may require networking over only tens of meters. Such distributed quantum computing can help to overcome individual quantum devices' hardware limitations. Quantum networking can also enhance quantum sensors through coordinated operation. (More details on this can be found in the Quantum Sensing section later in this chapter.)

Quantum networking entails many technical challenges, all focused on how to transfer quantum information between quantum computers or other quantum devices without loss. Because quantum information cannot be copied, it must be moved from point to point rather than replicated and then sent; this makes any losses during transmission extremely detrimental to quantum computers' operation.

One of these approaches depends on the computers involved being of the same design, thus eliminating the need to convert information from one form to another. This avoids the losses inherent in any such conversion. For example, to move information between quantum computers based on superconducting circuits, a networking design could require the qubits to remain in an ultra-low-temperature cryogenic environment. Such an environment would ensure that they maintain their quantum form and would minimize the risk of losses caused by noise, even very small vibrations or changes in temperature. This could be accomplished by housing the communicating computers within a single large cryogenic system or in a network of cryogenic environments connected by superconducting coaxial cables.

A second approach is transduction, which is the process of converting quantum information from one physical quantum system to another without losing its quantum properties.[23] Quantum transduction to optical photons (typically in fibers) is essential for networking quantum processors over long distances.[24] It is also essential whenever quantum information must be transmitted through room-temperature, non-vacuum environments.

Typical examples of transduction include converting quantum information from states of an atom or superconducting qubit to states of a photon for transmission over an optical fiber network. In these cases, the quantum transducer converts quantum signals encoded as atomic spins or microwave photons into optical photons and vice versa while preserving their quantum state.

For long-distance quantum communication (many kilometers and above), quantum repeaters may be needed. A quantum signal traveling long distances degrades because information-carrying photons are eventually lost during transmission due to absorption by the glass comprising the optical fiber.

To deal with this loss, quantum repeaters—essentially small quantum computers dedicated to a single function—must first build up entanglement between adjacent nodes of the network. They then use that entanglement along with classical communication to generate longer and longer range entanglement. Preserving the qubits and their entanglement in quantum memory at each node (i.e., at each repeater) as the chain grows, this process continues until the entanglement spans the entire network. At that point, the network can be employed—again in conjunction with classical communication—to teleport quantum information without errors or loss across its full extent.

Quantum networking entails many technical challenges, all focused on how to transfer quantum information between quantum computers or other quantum devices without loss.

Such repeaters, which are the subject of intense research activity, would make it possible to move quantum information from one place to another without making a copy of the original quantum state (which would otherwise violate the no-cloning theorem) in the presence of loss.

Key Developments

Some important technology developments in quantum communication in recent years include the following:

- **Interconnects** Researchers in China have built low-loss superconducting interconnects that enable high-fidelity quantum information transfer between modular quantum processor units.[25] A quantum interconnect is a device or system that links different quantum components—such as quantum processors, sensors, or memories—so they can exchange and process quantum information coherently.

- **Repeaters** A prototype quantum repeater that successfully distributes entanglement between two trapped-ion nodes separated by 50 kilometers (km) of optical fiber has been demonstrated.[26] This system integrates telecom-wavelength photon conversion to minimize transmission loss with quantum memories for storing entangled states and entanglement-swapping protocols, marking a critical advance toward scalable quantum networks.

- **Approaches to transduction** Some of the leading approaches to transduction include the following:

  - **Atomic decay–based transduction** This uses atoms or ions that absorb and emit light, helping transfer quantum information from one form (like microwave signals) to another (like optical light) without losing it.[27]

  - **Integrated microwave-to-optical transduction** Tiny optical circuits on chips convert quantum signals between microwave and optical light.[28] These devices can use either optomechanical or electro-optical interactions as the base physical process (see below). Integrated transducers can achieve high-rate transduction with reduced heat production.[29]

  - **Optomechanical transduction** Mechanical vibrations act as an intermediary between light and microwaves. These approaches have successfully demonstrated efficient bidirectional conversion[30] and microwave-optical entanglement.[31]

  - **Electro-optical transduction** The electric fields of the microwave photons modify the optical properties of a material to affect the microwave-to-optical conversion.[32]

  Despite significant progress, even the approaches that have succeeded in demonstrating microwave-optical entanglement currently do so at a quantum information transmission rate and fidelity significantly below what is needed to perform networked quantum computation. Importantly, current microwave-optical conversion rates are roughly six orders of magnitude smaller than the rate at which qubits in the same cryogenic environment can natively communicate with each other.

- **Transducer efficiency** New architectures have been developed that significantly increase the efficiency and speed of atom-to-photon quantum transducers.[33] Based on arrays of optical cavities, which are structures formed by pairs of highly reflective mirrors that can confine photons, the new architecture enables simultaneous transduction from many atoms. This advance is a step in the direction of higher-rate transmission in quantum networks.

- **Operational tests** Quantum-enabled data security affordances have been tested in operational scenarios by financial institutions such as HSBC,[34] JPMorgan Chase,[35] and Shinhan Bank in South Korea.[36] Other demonstrations have shown that QKD can operate over hundreds of kilometers[37] and even in space via the use of satellites and ground stations.[38]

Over the Horizon

Quantum communication is expensive and technologically difficult to put into place on a large scale. For it to be effective at such scale, it must solve problems less expensively than other approaches. As discussed in chapter 3, on cryptography and computer security, a variety of efforts are underway to deploy post-quantum encryption algorithms to resist quantum computing approaches to crack them. The security afforded by quantum communication will be valuable and important only if these efforts fail.

Even as the technology for quantum communication matures, classical networking will continue to handle most data traffic.[39] The vast majority of data in the world is classical, and classical networks are far more efficient and far faster for high-bandwidth communications. Quantum networks are inherently prone to loss and noise, limiting distance and speed, and their reliance on complex hardware like quantum repeaters restricts scalability. In short, quantum networks will at best complement, but not supersede, their classical counterparts in secure communication applications.

Nonetheless, it is expected that future networking between quantum data centers will enable joint, distributed quantum computations that surpass the capabilities of isolated quantum processors. This scaling approach is critical because of the physical and technical challenges in building large monolithic quantum computers. In this context, a "quantum internet" is likely to emerge.[40] This will likely involve localized quantum data centers performing the heavy lifting and more modest quantum devices securely querying them and communicating with one another. Networking between these quantum data centers will enable even larger calculations.

Finally, large-scale quantum networks may enable novel applications beyond secure communication and sensing. Speculatively, these include the following:

- **Quantum-assisted location verification and encryption** A protocol called quantum geo-encryption could allow data to be decrypted only at a specific geographic location or time, enhancing security against unauthorized or improper access.[41]

- **Entanglement-enhanced clock comparisons** Networks of optical clocks can surpass classical precision limits in frequency and time measurement, enabling applications in fundamental physics tests, navigation, and geodesy.[42]

- **Spoofing-proof timing synchronization** Using quantum entanglement properties for time synchronization could make timing systems much more resistant to spoofing or jamming attacks. This would improve the reliability of global navigation satellite systems, such as GPS, Russia's GLONASS, and the European Union's Galileo.[43]

Quantum Sensing

Essential Points

- Quantum sensing is the most mature of quantum technologies and is uniquely well suited for applications that involve small signals or that are delicate. These include astronomy (which requires capture of dim images), bioimaging (which requires that the light source not damage delicate specimens), and ultra-low-power platforms. Quantum sensing demonstrably excels in areas where classical probes are impractical, invasive, or inadequate, including gravitational-wave detection, precision timekeeping, and nanoscale field sensing.

- Quantum sensors increase their sensitivity by first suppressing classical sources of noise (e.g., technical or engineering noise). They then use quantum techniques for controlling quantum noise resulting from quantum effects such as the uncertainty principle. Quantum sensors do not eliminate all noise issues.

- True quantum advantage is not demonstrated by individual devices that show exceptional sensitivity only under idealized, highly controlled conditions. Rather it is demonstrated by a full system that outperforms a well-optimized classical baseline under equal resource constraints (size, weight, power, integration time). The system's sensitivity must reach and be maintained at the quantum noise limit in a robust, practically engineered package that doesn't require complex calibration and is rugged enough to use under changing environmental conditions.

- The advancement of practical quantum sensing is dominated by engineering challenges. These include the effective integration of system components, minimization of signal loss, and the development of compact, reliable photonic devices. Additionally, if sensors are networked, the challenges also include the establishment of precise timing and phase coordination, the reduction of losses in interconnections, and reliable synchronization across the network. Classical networking of quantum-enhanced nodes is likely to mature before fully entangled networks are fieldable.

- Networked and coordinated quantum sensing offer a number of benefits in principle, such as better ways to minimize the impact of noise and more accurate estimation of certain quantum wave properties important for sensing. However, realizing them in practice is a significant challenge separate from single-node engineering.

About Quantum Sensing

Sensing gathers information about the world, from telling time to detecting faint signals, such as light or gravitational fields. Many sensor improvements over the years have been driven by advances in materials, electronics, and data processing. However, as sensors are made more sensitive, they eventually hit the limits of classical measurement precision. These limits can sometimes be overcome by devoting more energy or time to the sensing process. Quantum sensing exploits quantum mechanics to achieve greater sensitivity per photon, per atom, per second, or per joule of energy. It does this through techniques such as entanglement, squeezed states of light, and quantum feedback.

- **Entanglement** refers to linking particles so their states become correlated, enhancing measurement precision.

- **Squeezed light** refers to light prepared in a way that reduces the noise due to random photon arrival times, thereby improving sensitivity beyond what is possible with classical instruments.

- **Quantum feedback** involves controlling quantum systems in real time to correct errors and boost measurement accuracy.

Quantum sensors measure diverse phenomena including time, magnetic and electric fields, mass, and forces such as gravity. To achieve their advanced sensitivity, these devices must protect the fragile quantum properties of their probes from the external environment through to the final measurement.

- For sensors based on matter—such as neutral atoms, ions, or nitrogen-vacancy (NV) centers in diamonds—this typically means shielding them to extend their coherence time (i.e., the time over which their probes retain their quantum state information).

- For photonic sensors, the dominant challenge is different: Optical loss must be minimized as light travels and interacts with the sample because lost photons irrevocably degrade the quantum advantage of the sensor. Furthermore, the final measurement itself must be nearly perfect; this often requires specialized hardware like high-quantum-efficiency photodetectors to ensure the delicate quantum information is read out without being destroyed.

Quantum sensing repurposes a fundamental challenge from quantum computing: While a quantum computer must be shielded from environmental noise, a sensor uses its quantum components as sensitive probes that intentionally interact with specific parts or properties of the environment to obtain information about it.

This distinction shows how the same quantum systems can be used differently by leveraging their unique properties for either computation or sensing. Similar observations apply for supporting technologies shared with quantum computing, such as lasers, photonics, and cryogenics. Overall, the field benefits from the fact that progress in one quantum area drives advances in others.

Key Developments

The applications in which quantum sensors are used today include scientific instrumentation, navigation, energy prospecting, biological imaging, and defense, among others. To support these and other applications, a number of underlying quantum sensing technologies have been developed. These include technologies for sensing magnetic and gravitational fields, quantum networking for sensors, and technologies related to operational integration. We address each in turn.

APPLICATION DOMAINS

Scientific instrumentation An example is the Laser Interferometer Gravitational-Wave Observatory (LIGO), a large-scale physics experiment designed to detect gravitational waves produced by cosmic events such as black holes and neutron star collisions. LIGO uses laser interferometry enhanced by squeezed light to measure changes that are far smaller than the width of a proton in the 4 km length of its arms,[44] which would not be possible with classical instruments. This quantum upgrade has nearly doubled the volume of the universe LIGO can observe. This means it can detect far more events, pushing the boundaries of our cosmic understanding.

Navigation Quantum sensors can facilitate GPS-independent navigation in two ways. First, they can provide superior inertial navigation through more accurate measurements of acceleration, rotation, and reduced sensor drift.[45] (Sensor drift is the gradual and undesirable change in a sensor's output over time, even when the input being measured remains constant. This results in a discrepancy between the sensor's readings and the true physical value of what is being measured.)

Second, quantum sensors can also support navigation through map matching. This involves sensors measuring subtle variations in the local electric or magnetic field as a function of position and aligning those measurements with a previously constructed map of the area to pinpoint their location.[46] The accuracy of this method depends on the availability of detailed field maps. In one demonstration, a quantum magnetic-anomaly navigation system was tested on aircraft and ground vehicles, achieving positioning accuracies on par with, or exceeding, GPS in some scenarios. This includes one case where the position accuracy was better than twenty-two meters, outperforming traditional inertial navigation systems by up to forty-six-fold.[47]

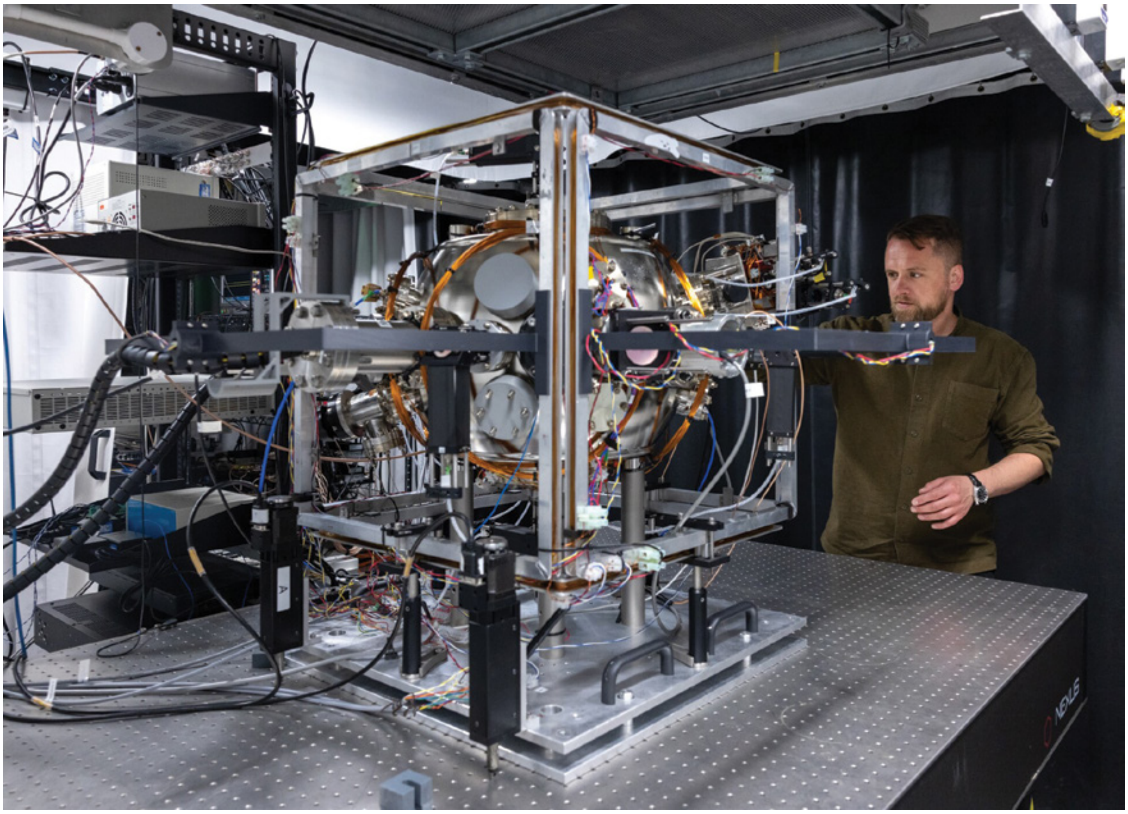
Energy and natural mineral prospecting In energy exploration,[48] quantum gravimeters can detect minute variations in Earth's gravitational field that provide information on the density and spatial structure of aquifers and hydrocarbon or minerals reserves. Quantum magnetometers can detect buried infrastructure, unexploded ordnance, and mineral deposits in real time, supporting safer and more efficient resource extraction.

Defense and national security Quantum sensing promises to support a wide range of battlefield applications, though none of those cited below have yet been matured fully into operational systems suitable for routine use. Navigation in GPS-denied environments is a critical need for mobile platforms and precision weapons, for which quantum accelerometers and gyroscopes are helpful (see figure 7.4). Quantum sensors for electromagnetic radiation detection and accurate timekeeping improve radar and electronic warfare systems by enhancing detection or improving resistance to jamming. Quantum-enhanced imaging methods can improve intelligence, surveillance, and reconnaissance capabilities, especially in low-visibility or degraded visual environments such as camouflage or dense foliage.

Quantum gravitometers may be able to assist in the more precise and certain locations of tunnels and underground bunkers.

Biological imaging and sensing Quantum sensors enable high-resolution optical imaging at ultra-low-light levels, minimizing light damage to sensitive samples. This allows researchers to observe living cells and tissues in their native states and supports noninvasive, high-precision, low-light imaging for biomedical research.[49] Quantum techniques can also extend these benefits to electron microscopy, enabling minimally destructive, nanoscale imaging of fragile specimens.

FIGURE 7.4 A quantum accelerometer from Imperial College London uses ultracold atoms to make highly accurate measurements



Source: Royal Navy

Quantum sensors can detect extremely small signals previously inaccessible to classical methods. For example, sensors based on NV centers in diamonds have measured microscopic magnetic and electric fields in single molecules and neurons and have also mapped heat production within cells.[50] Unlike traditional electrodes, quantum sensors probe living systems noninvasively, offering high spatial resolution and sensitivity. This enables precise study of brain function, neural communication, and neurological diseases.

A novel development is the use of an enhanced yellow fluorescent protein as a qubit.[51] This protein-based quantum sensor can be placed inside living systems to measure magnetic and electric fields, chemical changes, and protein interactions at the nanoscale. Because fluorescent proteins are naturally biocompatible, this approach allows quantum sensing within complex cellular environments. In essence, the biological qubit is like a temperature probe for the quantum age. However, instead of being sensitive to heat, it is sensitive to a suite of biologically relevant signals inside living systems.

SENSOR DEVELOPMENTS

The following are a sampling of some important developments in quantum sensors:

Cold-atom interferometers These use clouds of atoms cooled to near absolute zero to measure acceleration, rotation, and gravity with extraordinary precision. For example, cold-atom inertial sensors based on light-pulse atom interferometry have achieved sensitivity and accuracy levels that compete with—and sometimes surpass—traditional inertial sensors. Laboratory experiments have demonstrated high-precision measurements of acceleration and rotation, with technique improvements such as interleaved atom interferometry enabling measurements of rotation rate with high resolution and accuracy.[52] Chinese researchers successfully demonstrated that a cold-atom gyroscope has lower drift compared to traditional fiber-optic gyroscopes.[53]

NV center diamond sensors These are particularly useful for sensing magnetic fields.[54] They take advantage of a nitrogen atom next to a vacant site in a diamond's crystal structure in ways that enable detection of magnetic fields at sensitivities and resolutions far better than classical magnetic sensors. These sensors are thus well suited for the detection of weak, localized, or rapidly varying magnetic fields in domains such as biomedical imaging, scientific instrumentation more broadly, materials science, navigation, and defense.

Quantum radar Classical radar detects objects by sending microwave signals to a target and measuring what is reflected back. Quantum radar does the same using pairs of entangled photons, more often at optical frequencies rather than at microwave ones.[55] On theoretical grounds, this approach potentially offers modestly higher sensitivity in very noisy environments by leveraging quantum correlations rather than raw signal power. However, the increased complexity of the transmitters and receivers, and the modest enhancements promised, means that, in practice, a true quantum advantage in radar will remain elusive until significant photonic hardware challenges are overcome.

Networking for coordinated sensing Though networking for coordinated sensing has not been deployed in the field, even on an experimental basis, there is every reason to believe that quantum communication can enhance quantum sensors through coordinating their operation. This would improve sensitivity and precision beyond classical sensor arrays or individual quantum sensors. Applications span fields such as navigation, timing, environmental monitoring, geophysics, and medical imaging. As one example of recent theoretical work, researchers have shown how networking quantum sensors together can enable the highly selective detection of electromagnetic waves.[56] These networks can be tuned to ignore unwanted waves from specific directions while remaining sensitive to the desired signal.

> **In the coming decade, quantum sensing is expected to significantly impact fields such as biology, medicine, navigation, and geoscience.**

Operational integration Once the foundational science of a particular sensor is validated in the laboratory, its transition to field deployment introduces substantial engineering challenges. This is because performance that is robust under controlled lab conditions often declines under real-world stressors. Recognizing these issues, the Defense Advanced Research Projects Agency (DARPA) has initiated programs that tackle this operational challenge from different angles. Its Robust Quantum Sensors program, for example, focuses on system-level ruggedization. It aims to make entire complex sensor systems (like cold-atom interferometers) resilient enough to maintain high performance while mounted on moving military platforms.

A complementary, component-level approach is taken by DARPA's Intensity-Squeezed Photonic Integration for Revolutionary Detectors program.[57] This effort focuses on miniaturization and integration. It seeks to take the sophisticated squeezed-light technology used in massive experiments like LIGO and shrink it to a compact photonic chip.[58] When it happens in any given use case, the shift from large, stationary experimental setups to compact, rugged field sensors is a significant milestone, enabling systems that are valuable to military end users.

Although not funded by DARPA, one example of such work is the demonstration of the integration and miniaturization of accelerometers based on cold-atom interferometry. This is done by combining traditionally separate components into a compact unit suitable for field deployment on operational platforms. This innovation reduces the size and complexity from large optical-table systems to a rugged package roughly the size of a shoebox.[59]

Over the Horizon

Especially exciting now is the development of new applications of quantum sensing technologies. These include portable brain scanners that use atomic magnetometers; quantum-enhanced microscopes that can look inside living cells without damaging them; and gravimeters that can detect underground cavities without digging.

In the coming decade, quantum sensing is expected to significantly impact fields such as biology, medicine, navigation, and geoscience. In the life sciences, emerging protein-based sensors may eventually enable measurements of signals inside living cells, opening new opportunities in neuroscience and bioanalytics. For example, quantum sensors may enable in-vivo imaging through enhanced microspectroscopy.[60] Theoretical work also suggests that entangled photon imaging could further improve the imaging of tissues and biological organisms, although this remains largely conceptual.[61]

In navigation and geoscience, quantum gravimeters, gyroscopes, and portable clocks could strengthen navigation where GPS is unavailable while improving climate monitoring and resource mapping. Progress in integrated quantum photonics is also anticipated to deliver substantial gains in power efficiency, bringing new sensing capabilities to power-constrained platforms.

At the same time, a new frontier lies in algorithmic sensing: Rather than treating sensors as passive data collectors, quantum algorithms could directly use quantum sensors to gather and process data. Early theoretical work points to the possibility of extracting extremely weak signals from noise or scanning large, uncertain parameter spaces far more efficiently than classical approaches. If realized, such algorithmic strategies would link quantum sensors and quantum computers in real-time feedback loops, expanding capabilities in areas from spectrum monitoring to distributed sensor networks.

Next-generation quantum sensors promise to extend the march toward lower drift, lower power consumption, and higher precision. Therefore, it seems reasonable to expect that these techniques, combined with large-scale classical data-fusion techniques, will continue to progress in such detection tasks. In short, we anticipate continued incremental gains in sensor capabilities as the result of continued progress in quantum sensing, rather than a large, discrete jump in overall performance.

Policy Issues

The American Advantage

America currently leads the world in many of the most advanced quantum computing technologies, including superconducting circuits, neutral atoms, and trapped ions. This leadership is the result of the following:

- A stable pipeline of experimentally demonstrated approaches, techniques, and materials emerging from basic research labs on university campuses
- A healthy ecosystem of start-ups and large companies exploring commercialization opportunities
- A forward-looking, entrepreneurial mindset that has been willing to support the pursuit of new opportunities before their payoff is clear

The United States doesn't necessarily lead in developing enabling technologies, such as electronics, optics, and cryogenics. Nor is it the fastest at commercializing or scaling up quantum devices; furthermore, its skilled and unskilled labor is relatively expensive. Nevertheless, the national quantum ecosystem that's been built since World War II to nurture and develop novel creative ideas and support their commercialization has given America a significant competitive advantage.

Other countries are now investing significant amounts of public capital into quantum technologies—in particular, China, which has made the domain one of the main technology priorities in its five-year plans.[62] To sustain America's advantage in quantum technologies, several areas of policy are important.

> **The national quantum ecosystem that's been built since World War II . . . has given America a significant competitive advantage.**

Support for Basic Science Research

In the past five years, both start-ups and tech giants have increasingly taken on the task of building large-scale, commercial quantum computing. Industrial efforts that focus on a wide variety of qubit technologies and their associated architectures now exist. These efforts typically focus on building toward the million-physical-qubit systems necessary for useful error-corrected computation. Thus, they emphasize the scaling of existing technologies, with relatively small tweaks to hardware and algorithms.

By contrast, academic efforts tend to focus on discovery of entirely new qubit platforms, improving qubit properties in existing platforms, creating new routes to scaling, and developing new quantum algorithms. Unlike many fields where academic efforts are disconnected from industrial realities, this high-risk basic academic science remains crucial to the continued progress of private-sector efforts. Ideas from academia are regularly adopted by both large companies and start-ups.

The Quantum Workforce

On university campuses—where the basic science powering quantum innovations occurs—it is critically important to ensure access to doctoral students from abroad and particularly from China. There is an insufficient supply of adequately skilled US-trained undergraduates. Therefore, cutting the flow of international students actively hinders American competitiveness in quantum technology. In almost every university research laboratory, graduate students, as part of their training, provide much of the day-to-day labor necessary for advancing the state-of-the-art science.[63] Furthermore, after these international students have graduated, they disproportionately move into the quantum workforce in America; ensuring that this trend remains possible is essential to the current and future dynamism of America's quantum creativity engine.

Supply Chain

Quantum hardware requires an increasingly diverse array of support technologies. On the materials side, this includes electro-optic and acousto-optic crystals, high purity aluminum, and rare earth metals. Among assembled technologies, it includes lasers, optical modulators, single photon detectors, and cryogenic refrigerators. Many of these technologies are produced abroad, typically at a cost many times lower than that of US manufacturers. In some cases, no US alternatives are available. Because quantum technology remains largely in a research and development phase, tariffs on these support technologies cannot be passed on to consumers, stifling the ability of quantum scientists and engineers to lead the way in developing them.

Competition with China

There is currently an active competition between the United States and China in quantum technologies. China is leading in overall investment, at around $15 billion,[64] while the United States lags behind at around $8 billion ($4 billion of private investment plus $4 billion of public funding).[65]

As of August 2025, the largest controlled system of neutral atoms was produced in China,[66] as was the largest system of trapped ions.[67] Similarly, Chinese demonstrations of ground-to-satellite quantum networking via the Micius satellite and a 700-fiber ground-based communications network exceed (at least in terms of scale) anything yet attempted in the United States.[68] In all of these cases, the techniques have been pushed to their limits, leveraging impressive integration of technologies beyond the traditional quantum ecosystem. Regardless of whether these technologies ultimately prove commercially useful, efforts to develop them have unquestionably contributed to further growth of the Chinese talent base in quantum science and technology.

By contrast, the first demonstrations of quantum logic in neutral atom arrays, superconducting circuits, and ion traps all originated from the United States. Most new approaches and technologies continue to emerge from the US academic ecosystem. DARPA programs in quantum computing, networking, and sensing sharpen and focus these efforts, and a broad ecosystem of basic science funded through the National Science Foundation, US Department of Energy, and US Department of Defense support this innovation engine. Examples of US government funding for basic research in quantum science and technology include the National Quantum Initiative, which was signed into law in early 2019,[69] and the Quantum Benchmarking Initiative, launched in 2024 by DARPA.[70]

Assuming that there is sustained support for basic research in quantum technologies, it seems likely that US innovation will continue to drive progress and that China will follow in ideas but lead in scaling. The gap, however, is closing, as China recruits more talent and grows its basic science portfolio.

NOTES

1. "Nobel Prize in Physics 2025," NobelPrize.org, Nobel Prize Outreach 2025, October 20, 2025, https://www.nobelprize.org/prizes/physics/2025/summary/.

2. Zack Savitsky, "Debate Erupts Around Microsoft's Blockbuster Quantum Computing Claims," *Science* 387(6741): 1338–39, March 28, 2025, https://www.science.org/content/article/debate-erupts-around-microsoft-s-blockbuster-quantum-computing-claims.

3. David S. Wang, Austin G. Fowler, and Lloyd C. L. Hollenberg, "Surface Code Quantum Computing with Error Rates over 1%," *Physical Review A* 83, no. 2 (2011): 020302, https://doi.org/10.1103/PhysRevA.83.020302.

4. An analogy to classical computing would be that it is possible to build digital computers out of transistors and of vacuum tubes. For a variety of reasons (e.g., cost, energy consumption), transistor technology proved decisively superior as the fundamental hardware building block of classical computers.

5. Alkim Bozkurt, Han Zhao, Chaitali Joshi, Henry G. LeDuc, Peter K. Day, and Mohammad Mirhosseini, "A Quantum Electromechanical Interface for Long-Lived Phonons," *Nature Physics* 19, no. 9 (2023): 1326–32, https://doi.org/10.1038/s41567-023-02080-w.

6. Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, et al., "Logical Quantum Processor Based on Reconfigurable Atom Arrays," *Nature* 626 (2024): 58–65, https://doi.org/10.1038/s41586-023-06927-3.

7. Craig Gidney, "How to Factor 2048 Bit RSA Integers with Less than a Million Noisy Qubits," arXiv:2505.15917, version 1, preprint, arXiv, May 21, 2025, https://doi.org/10.48550/arXiv.2505.15917.

8. Michael E. Beverland, Prakash Murali, Matthias Troyer, et al., "Assessing Requirements to Scale to Practical Quantum Advantage," arXiv:2211.07629, preprint, arXiv, November 14, 2022, https://doi.org/10.48550/arXiv.2211.07629.

9. Matt Swayne, "Quantum Computing Roadmaps and Predictions of Leading Players," *The Quantum Insider*, May 16, 2025, https://thequantuminsider.com/2025/05/16/quantum-computing-roadmaps-a-look-at-the-maps-and-predictions-of-major-quantum-players/.

10. Hari P. Paudel, Madhava Syamlal, Scott E. Crawford, et al., "Quantum Computing and Simulations for Energy Applications: Review and Perspective," *ACS Engineering Au* 2, no. 3 (2022): 151–96, https://doi.org/10.1021/acsengineeringau.1c00033.

11. Mauro D'Arcangelo et al., "Leveraging Analog Quantum Computing with Neutral Atoms for Solvent Configuration Prediction in Drug Discovery," *Physical Review Research* 6, no. 4 (2024), https://doi.org/10.1103/PhysRevResearch.6.043020.

12. Ehud Altman et al., "Quantum Simulators: Architectures and Opportunities," *PRX Quantum* 2, no. 1 (2021): 017003; A. J. Daley, I. Bloch, C. Kokail, et al., "Practical Quantum Advantage in Quantum Simulation," *Nature* 607 (2022): 667–76.

13. Craig Gidney, "How to Factor 2048 Bit RSA Integers with Less than a Million Noisy Qubits." arXiv:2505.15917, version 1, preprint, arXiv, May 21, 2025, https://doi.org/10.48550/arXiv.2505.15917. The five-to-fifteen-year time frame comes from "DARPA Eyes Companies Targeting Industrially Useful Quantum Computers," DARPA, April 3, 2025, https://www.darpa.mil/news/2025/companies-targeting-quantum-computers.

14. Oskar Leimkuhler and K. Birgitta Whaley, "Exponential Quantum Speedups in Quantum Chemistry with Linear Depth," arXiv:2503.21041, preprint, arXiv, May 16, 2025, https://doi.org/10.48550/arXiv.2503.21041.

15. Benedikt Fauseweh, "Quantum Many-Body Simulations on Digital Quantum Computers: State-of-the-Art and Future Challenges," *Nature Communications* 15, no. 1 (2024): 2123, https://doi.org/10.1038/s41467-024-46402-9.

16. Hsin-Yuan Huang et al., "Power of Data in Quantum Machine Learning," *Nature Communications* 12, no. 1 (2021): 2631.

17. In general, a polynomial speedup is useful for turning a problem that is already doable in a plausible amount of time into one that is doable in a meaningfully shorter amount of time. Instead of a billion years turning into a few minutes, polynomial speedups turn days or weeks of computation into minutes and can be fairly regarded as highly useful but not transformational.

18. See, for example, Seunghoon Lee, Joonho Lee, Huanchen Zhai, et al., "Evaluating the Evidence for Exponential Quantum Advantage in Ground-State Quantum Chemistry," *Nature Communications* 14, no. 1 (2023): 1952, https://doi.org/10.1038/s41467-023-37587-6.

19. Yulin Wu et al., "Strong Quantum Computational Advantage Using a Superconducting Quantum Processor," *Physical Review Letters* 127, no. 18 (2021), https://doi.org/10.1103/PhysRevLett.127.180501.

20. William K. Wootters and Wojciech H. Zurek, "A Single Quantum Cannot Be Cloned," *Nature* 299 (1982): 802–3.

21. "Quantum Safe Cryptography and Security," ETSI White Paper #8, European Telecommunications Standards Institute, June 2015, https://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf.

22. Sunil K. Singh, Sudhakar Kumar, Anureet Chhabra, et al., "Advancements in Secure Quantum Communication and Robust Key Distribution Techniques for Cybersecurity Applications," *Cyber Security and Applications* 3 (December 2025): 100089, https://doi.org/10.1016/j.csa.2025.100089.

23. David Awschalom et al., “Development of Quantum Interconnects (Quics) for Next-Generation Information Technologies,” *Prx Quantum* 2 (2021): 017002.

24. Marcello Caleffi, Laura d’Avossa, Xu Han, and Angela Sara Cacciapuoti, “Quantum Transduction: Enabling Quantum Networking,” arXiv:2505.02057, preprint, arXiv, May 4, 2025, https://doi.org/10.48550/arXiv.2505.02057.

25. Simon Storz, Josua Schär, Anatoly Kulikov, et al., “Loophole-Free Bell Inequality Violation with Superconducting Circuits,” *Nature* 617, no. 7960 (2023): 265–70, https://doi.org/10.1038/s41586-023-05885-0.

26. V. Krutyanskiy, M. Canteri, M. Meraner, et al., “Telecom-Wavelength Quantum Repeater Node Based on a Trapped-Ion Processor,” *Physical Review Letters* 130, no. 21 (2023): 213601, https://doi.org/10.1103/PhysRevLett.130.213601.

27. Aishwarya Kumar, Aziza Suleymanzade, Mark Stone, et al., “Quantum-Enabled Millimetre Wave to Optical Transduction Using Neutral Atoms,” *Nature* 615, no. 7953 (2023): 614–19, https://doi.org/10.1038/s41586-023-05740-2.

28. Nikolai Lauk, Neil Sinclair, Shabir Barzanjeh, et al., “Perspectives on Quantum Transduction,” *Quantum Science and Technology* 5, no. 2 (2020): 020501, https://doi.org/10.1088/2058-9565/ab788a.

29. Galan Moody, Volker J. Sorger, Daniel J Blumenthal, et al., “2022 Roadmap on Integrated Quantum Photonics,” *Journal of Physics: Photonics* 4, no. 1 (2022): 012501, https://doi.org/10.1088/2515-7647/ac1ef4.

30. R. W. Andrews, R. W. Peterson, T. P. Purdy, et al., “Bidirectional and Efficient Conversion Between Microwave and Optical Light,” *Nature Physics* 10, no. 4 (2014): 321–26, https://doi.org/10.1038/nphys2911.

31. Srujan Meesala, David Lake, Steven Wood, et al., “Quantum Entanglement Between Optical and Microwave Photonic Qubits,” *Physical Review X* 14, no. 3 (2024), https://doi.org/10.1103/PhysRevX.14.031055.

32. Charles Möhl, Annina Riedhauser, Max Glantschnig, et al., “Bidirectional Microwave-Optical Conversion with an Integrated Soft-Ferroelectric Barium Titanate Transducer,” preprint, January 26, 2025, https://doi.org/10.1103/1gvf-w6lx.

33. For theoretical work on the superiority of cavity arrays to lenses, see Josiah Sinclair, Joshua Ramette, Brandon Grinkemeyer, Dolev Bluvstein, Mikhail Lukin, and Vladan Vuletić, “Fault-Tolerant Optical Interconnects for Neutral-Atom Arrays,” arXiv:2408.08955, preprint, arXiv, August 16, 2024, https://doi.org/10.48550/arXiv.2408.08955. For experimental work implementing cavity arrays, see Adam L. Shaw, Anna Soper, Danial Shadmany, et al., “A Cavity Array Microscope for Parallel Single-Atom Interfacing,” arXiv:2506.10919, preprint, arXiv, June 12, 2025, https://doi.org/10.48550/arXiv.2506.10919.

34. “HSBC Pioneers Quantum Protection for AI-Powered FX Trading,” news release, HSBC, December 23, 2023, https://www.hsbc.com/news-and-views/news/media-releases/2023/hsbc-pioneers-quantum-protection-for-ai-powered-fx-trading.

35. “JPMorgan Chase to Implement a Quantum-Secured Network,” *Quantum Herald*, n.d., accessed September 27, 2025, https://www.quantumsecuritydefence.com/news/jpmorgan-chase-to-implement-a-quantum-secured-network.

36. “Toshiba Digital Solutions and KT Demonstrate Hybrid Quantum Secure Communications with South Korea’s Shinhan Bank,” Toshiba.com, April 19, 2024, https://www.global.toshiba/ww/news/digitalsolution/2024/04/news-20240419-01.html.

37. Shi-Chang Zhuang, Bo Li, Ming-Yang Zheng, et al., “Ultrabright-Entanglement-Based Quantum Key Distribution over a 404-Km-Long Optical,” arXiv:2408.04361, version 1, preprint, arXiv, August 8, 2024, https://doi.org/10.48550/arXiv.2408.04361.

38. Juan Yin, Yu-Huai Li, Sheng-Kai Liao, et al., “Entanglement-Based Secure Quantum Cryptography over 1,120 Kilometres,” *Nature* 582, no. 7813 (2020): 501–5, https://doi.org/10.1038/s41586-020-2401-y.

39. Joseph M. Lukens, Nicholas A. Peters, and Bing Qi, “Hybrid Classical-Quantum Communication Networks,” arXiv:2502.07298, preprint, arXiv, July 24, 2025, https://doi.org/10.48550/arXiv.2502.07298.

40. H. Jeff Kimble, “The Quantum Internet,” *Nature* 453.7198 (2008): 1023–30.

41. Robert Malaney, “Quantum Geo-Encryption,” 2016 IEEE Global Communications Conference (IEEE Press, 2016), 1–6, https://doi.org/10.1109/GLOCOM.2016.7842191.

42. “An Elementary Network of Entangled Optical Atomic Clocks?,” University of Oregon, June 2, 2023, https://omq.uoregon.edu/events/elementary-network-entangled-optical-atomic-clocks.

43. Abyad Enan, Mashrur Chowdhury, Sagar Dasgupta, and Mizanur Rahman, “Quantum-Classical Hybrid Framework for Zero-Day Time-Push GNSS Spoofing Detection,” arXiv:2508.18085, preprint, arXiv, August 25, 2025, https://doi.org/10.48550/arXiv.2508.18085.

44. M. Tse et al., “Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy,” *Physical Review Letters* 123, no. 23 (2019), https://doi.org/10.1103/PhysRevLett.123.231107.

45. Mary Burkey, *How Quantum Sensing Will Solve GPS Denial in Warfare*, Lawrence Livermore National Laboratory Center for Global Security Research Fellows Publication, June 2025, https://cgsr.llnl.gov/sites/cgsr/files/2025-06/Burkey_QS_final.pdf.

46. Samuel Lellouch and Michael Holynski, “Integration of a High-Fidelity Model of Quantum Sensors with a Map-Matching Filter for Quantum-Enhanced Navigation,” *Quantum Science and Technology* 10, no. 4 (2025): 045007, https://doi.org/10.1088/2058-9565/adf2d9.

47. Murat Muradoglu, Mattias T. Johnsson, Nathanial M. Wilson, et al., “Quantum-Assured Magnetic Navigation Achieves Positioning Accuracy Better than a Strategic-Grade INS in Airborne and Ground-Based Field Trials,” arXiv:2504.08167, preprint, arXiv, April 10, 2025, https://doi.org/10.48550/arXiv.2504.08167.

48. Scott E. Crawford, Roman A. Shugayev, Hari P. Paudel, et al., “Quantum Sensing for Energy Applications: Review and Perspective,” *Advanced Quantum Technologies* 4, no. 8 (2021), https://doi.org/10.1002/qute.202100049.

49. Hoda Lotfipour, Hassan Sobhani, Mohamad Taghi Dejpasand, and Morteza Sasani Ghamsari, “Application of Quantum Imaging in Biology,” *Biomedical Optics Express* 16, no. 8 (2025): 3349–77, https://doi.org/10.1364/BOE.566801.

50. Nikolaj Hansen et al., “Microscopic-Scale Magnetic Recording of Brain Neuronal Electrical Activity Using a Diamond Quantum Sensor,” *Sci. Rep* 13, no. 1 (2023); Nabeel Aslam et al., “Quantum Sensors for Biomedical Applications,” *Nature Review Physics* 5, no. 3 (2023).

51. Jacob S. Feder, Benjamin S. Soloway, Shreya Verma, et al., “A Fluorescent-Protein Spin Qubit,” *Nature* 645, no. 8079 (2025): 73–9, https://doi.org/10.1038/s41586-025-09417-w.

52. Remi Geiger, Arnaud Landragin, Sébastien Merlet, and Franck Pereira Dos Santos, "High-Accuracy Inertial Measurements with Cold-Atom Sensors," *AVS Quantum Science* 2, no. 2 (2020): 024702, https://doi.org/10.1116/5.0009093.

53. Jinting Li, Xi Chen, Danfang Zhang, et al., "Realization of a Cold Atom Gyroscope in Space," *National Science Review* 12, no. 4 (2025): nwaf012, https://doi.org/10.1093/nsr/nwaf012.

54. Ning Wang and Jianming Cai, "Hybrid Quantum Sensing in Diamond," *Frontiers in Physics* 12 (February 2024), https://doi.org/10.3389/fphy.2024.1320108.

55. Athena Karsa, Alasdair Fletcher, Gaetana Spedalieri, and Stefano Pirandola, "Quantum Illumination and Quantum Radar: A Brief Overview," *Reports on Progress in Physics* 87, no. 9 (2024): 094001, https://doi.org/10.1088/1361-6633/ad6279.

56. Arne Hamann, Paul Aigner, Pavel Sekatski, and Wolfgang Dür, "Selective and Noise-Resilient Wave Estimation with Quantum Sensor Networks," *Quantum Science and Technology* 10, no. 3 (2025): 035028, https://doi.org/10.1088/2058-9565/add61b.

57. US Defense Advanced Research Projects Agency, "Intensity-Squeezed Photonic Integration for Revolutionary Detectors (INSPIRED)," accessed October 20, 2025, https://www.darpa.mil/research/programs/intensity-squeezed-photonic.

58. Jon Harper, "DARPA Eyeing New Quantum Sensing Program," *DefenseScoop*, December 30, 2024, https://defensescoop.com/2024/12/30/darpa-eying-new-quantum-sensing-program-robust-quantum-sensors-roqs/.

59. Jongmin Lee, Roger Ding, Justin Christensen, et al., "A Compact Cold-Atom Interferometer with a High Data-Rate Grating Magneto-Optical Trap and a Photonic-Integrated-Circuit-Compatible Laser System," *Nature Communications* 13, no. 1 (2022): 5131, https://doi.org/10.1038/s41467-022-31410-4.

60. Tian Li, Vsevolod Chevurkanov, Vladislav Yakovlev, "Harnessing Quantum Light for Microscopic Biomechanical Imaging of Cells and Tissues," *PNAS* 121, no. 45 (2024).

61. Yide Zhang, Zhe He, Xin Tong, David Garrett, et. al., "Quantum Imaging of Biological Organisms Through Spatial and Polarization Entanglement," *Science Advances* 10, no. 10 (2024).

62. State Council, People's Republic of China, "China to Include Quantum Technology in Its 14th Five-Year Plan," CGTN News Service, October 22, 2020, https://english.www.gov.cn/news/videos/202010/22/content_WS5f90e700c6d0f7257693e3fe.html.

63. National Science Board, "Science & Engineering Indicators: Human Resources and Academic R&D" (NSB2023-26), accessed October 20, 2025, https://ncses.nsf.gov/pubs/nsb202326/human-resources-and-academic-r-d.

64. Brad Smith, "Investing in American Leadership in Quantum Technology," *Microsoft on the Issues*, April 28, 2025, https://blogs.microsoft.com/on-the-issues/2025/04/28/investing-in-american-leadership-quantum/.

65. "Special Gaps Analysis-Quantum," Special Competitive Studies Project, 2025, https://www.scsp.ai/reports/2025-gaps-analysis/gaps-analysis/quantum/.

66. Lin Rui et al., "AI-Enabled Parallel Assembly of Thousands of Defect-Free Neutral Atom Arrays," *Physical Review Letters* 135, no. 6 (2025): 060602.

67. Guo S-A. et al., "A Site-Resolved Two-Dimensional Quantum Simulator with Hundreds of Trapped Ions," *Nature* 630, no. 8017 (2024): 613–18.

68. Yu-Ao Chen, Qiang Zhang, Teng-Yun Chen, et al., "An Integrated Space-to-Ground Quantum Communication Network over 4,600 Kilometres," *Nature* 589, no. 7841 (2021): 214–19, https://doi.org/10.1038/s41586-020-03093-8.

69. Martin Giles, "President Trump Has Signed a $1.2 Billion Law to Boost US Quantum Tech," *MIT Technology Review*, December 22, 2018, https://www.technologyreview.com/2018/12/22/138149/president-trump-has-signed-a-12-billon-law-to-boost-us-quantum-tech/.

70. US Defense Advanced Research Projects Agency, "Quantum Benchmarking Initiative (QBI)," accessed October 20, 2025, https://www.darpa.mil/research/programs/quantum-benchmarking-initiative.

STANFORD EXPERT CONTRIBUTORS

Dr. Jon Simon
SETR Faculty Council, Joan Reinhart Professor of Physics, and Professor of Applied Physics

Dr. Amir Safavi Naeini
Associate Professor of Applied Physics and, by courtesy, of Electrical Engineering

Dr. David Schuster
Professor of Applied Physics

Ms. Junna Gui
SETR Fellow and Q-Farm Program Manager

08

ROBOTICS

KEY TAKEAWAYS

- Artificial intelligence holds significant potential to advance complex robotic systems, but the speed of future advances will depend on the availability of high-quality training data and the systematic integration of data-rich foundation models, simulated interactions between robots and their environment, and understanding of the real physical world.
- Humanoid robots show promise for specialized industrial and healthcare roles, although widespread adoption of them faces challenges linked to their cost, technical complexity, energy efficiency, safety, and training data quality.
- Advances in autonomous, low-cost, and communication-resilient robotic systems are transforming important aspects of modern warfare.

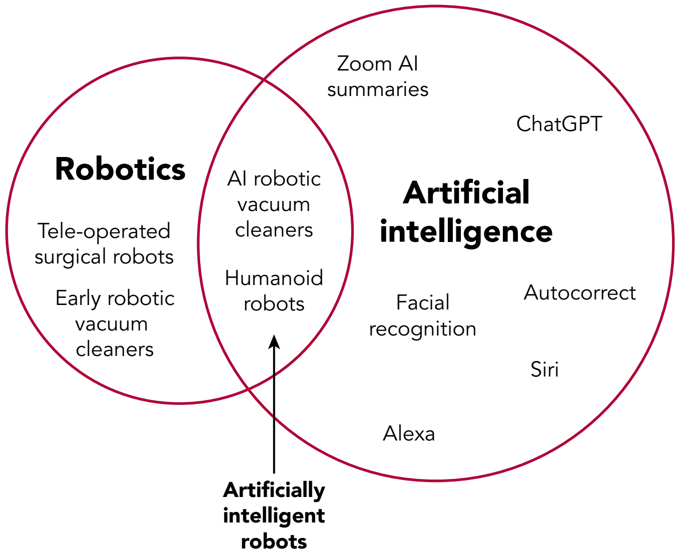
Overview

A robot is an engineered physical entity with ways of sensing itself or the world around it and of creating physical effects on that world.[1] Robots must integrate many different component technologies to combine perception of the environment with action in it. Perception requires generating representations of the robot's environment and its interaction with its surroundings. Action requires the robot to make physical changes to itself or the environment based on those perceptions.

The key engineering challenges in robotics involve the design of components, integration of these components within a robot's body, and algorithms that enable system-level functionality to allow a robot to perform intended tasks in different settings and environments. Important component technologies include:

- Actuators that enable movement, such as motors and grasping appendages.

FIGURE 8.1 Not all robots use artificial intelligence



- Sensors that receive real-time input about the immediate physical environment of the robot and the robot's own configuration.
- Control systems that decide what the robot should do based on sensor readings.
- Structural materials that robots are made of. Those built from rigid materials typically interact with their operating environments in highly prescribed and structured ways. "Soft" robots, which are flexible and conform to their surroundings, can offer better performance in more unstructured and chaotic environments.[2]
- Power sources that can be tethered to a robot or are untethered. A robot tethered to a "base station" can be energized from a power source on that base indefinitely, while untethered robots need self-contained power sources or sources that harvest energy from the environment.
- Real-time computing that determines the specific timeframes in which operations of robots take place ensures, for example, that a robotic arm in a workplace will stop very quickly if the robot detects a human in its immediate proximity.

Finally, some robots use computer vision and other types of artificial intelligence (AI) for understanding their environments and decision making, but robotics and AI do not always go together (see examples in figure 8.1). Robots with varying degrees of autonomy have been used in everything from delicate surgical procedures to space exploration.

Examples of robots include self-driving cars, drones, humanoids (i.e., a robot that mimics human form and motions), manipulators used in manufacturing and warehousing, and tele-operated surgical instruments. They can range in size from millimeter-scale soft medical devices that navigate vessels in the brain to large land vehicles and excavators for mining and construction.

The form factor of a robot—its overall size, shape, and physical layout—has far-reaching implications because it determines in large part what the robot can do, how well it can do it, and if and how people interact with it. A robot's form factor dictates how it moves, manipulates and senses things, and carries loads. It also influences the environments in which a robot can operate, its stability, and its capacity to withstand adverse conditions. Additionally, a robot's form factor affects its safety (including ease of use around people) and the cost of manufacturing it, its energy efficiency, and how easy or hard it is to repair and upgrade. Finally, a robot's size and appearance can also have significant regulatory and social implications, including how widely it is accepted by the public.

Key Developments

The development of robots is influenced by a complex mix of factors: technological advances in areas such as mechanical design, sensors, materials, actuators, AI, and control theory (i.e., the use of algorithms and feedback to manipulate the behavior of dynamical systems); their potential to solve specific tasks or problems; economic considerations including cost and market demand; safety and ethical regulations; global security concerns; and social acceptance shaped by cultural and human factors.

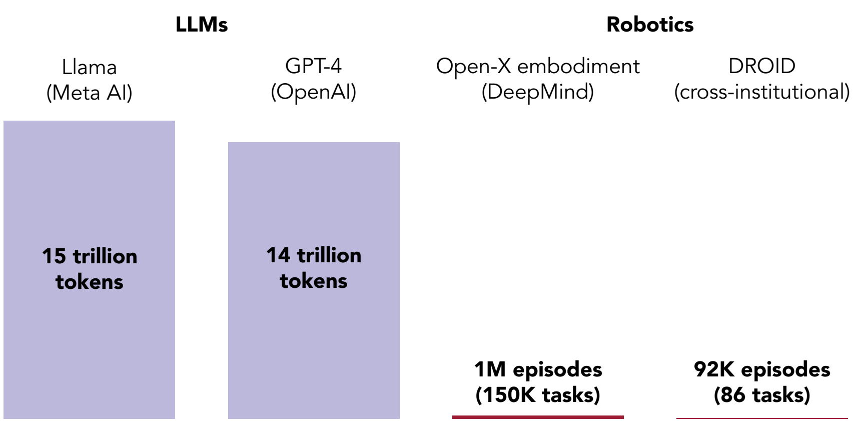
Some of the most important current developments in robotics are the role of data in AI for robotics, the development of humanoid robots, and the use of robotics in warfare.

The Role of Data in Artificial Intelligence for Robotics

The recent acceleration of AI, most notably through the creation of popular tools like ChatGPT, demonstrates how training AI models using large-scale data can drive remarkable technological advancements and economic gain. Robotics, however, faces unique challenges. Unlike digital text processing, which uses vast amounts of readily available textual data on the internet, robots require very detailed, specialized information, including visual data and sensor-based measurements of touch, precise motion, and machines' physical interactions with their surroundings.

Gathering such data at scale is both expensive and time-consuming, creating a significant bottleneck in the development of reliable, large-scale robot automation. As illustrated in figure 8.2, a large language model (LLM) today is typically trained on trillions

FIGURE 8.2 Tokens for LLMs versus episodes for robot models



Sources: Meta AI: "Introducing Meta Llama 3: The Most Capable Openly Available LLM to Date," Meta AI, April 18, 2024, https://ai.meta.com/blog/meta-llama-3/. Open AI: Julie Chang, host, *Tech News Briefing*, podcast, "The Internet May Be Too Small for the AI Boom, Researchers Say," WSJ Podcasts, April 15, 2024, https://www.wsj.com/podcasts/tech-news-briefing/the-internet-may-be-too-small-for-the-ai-boom-researchers-say/63680C0D-69FE-437C-98F8-B678DD1F7536. Industrial: Quan Vuong and Pannag Sanketi, "Scaling Up Learning Across Many Different Robot Types," Google DeepMind, October 3, 2023, https://deepmind.google/discover/blog/scaling-up-learning-across-many-different-robot-types/. Academic: Alexander Khazatsky, Karl Pertsch, Suraj Nair, et al., "DROID: A Large-Scale In-The-Wild Robot Manipulation Dataset," arXiv:2403.12945, preprint, arXiv, April 22, 2025, https://doi.org/10.48550/arXiv.2403.12945.

Viewed through a global lens, a race in robotic automation is already underway, driven primarily by the rapidly growing computational resources necessary to capture and analyze complex datasets.

of tokens, whereas the largest datasets to date for robot models amount to at most a million or so episodes. (Tokens are the basic units that LLMs read and generate to process text; they can be a whole word, part of a word, or punctuation. Episodes are the basic units that robot foundation models process—a full sequence of observations, actions, and outcomes during one task, capturing what the robot saw and did over time.)

Simulation is often offered as a cheaper and safer alternative to collecting real-world data at scale for robot automation. Unfortunately, simulations frequently fail to replicate the complexity and unpredictability of physical environments and naturally favor scenarios for which they have already been prepared. The result is that using a simulation yields data with less effort but also with less fidelity to real-world situations. Simulations are still valuable tools, but their usefulness and reliability ultimately depend on calibrating them with extensive real-world data that provides some measure of ground truth.

To address these challenges, a hybrid strategy of blending advanced AI methods with proven engineering approaches is often necessary. This approach ensures that robots are more robust and capable of handling uncertainty. Crucially, high-quality data in large quantities is vital for successful outcomes. In efforts to collect such data, disparities between different regions of the world—and differences in things such as rules surrounding use of personal data—need to be taken into account. A healthcare model trained on Chinese data might perform poorly in the European Union due to differences in healthcare systems and demographics, leading to unintended consequences.

Viewed through a global lens, a race in robotic automation is already underway, driven primarily by the rapidly growing computational resources necessary to capture and analyze complex datasets. Even the most well-resourced academic institutions in the United States often have access to significantly less computational power compared to the resources available in private industry. Nevertheless, industry continues to rely heavily on foundational research conducted in academia, highlighting the importance of sustaining robust academic capabilities to maintain technological leadership.

Humanoid Robots

Robots may be useful for improving the US manufacturing base, reducing supply chain vulnerabilities, delivering eldercare, enhancing food production, tackling the housing shortage, improving energy sustainability, and performing almost any task involving physical presence. One type of device—the humanoid—is promoted as a solution to labor shortages in industries such as logistics, manufacturing, and hospitality. For example, humanoid robots have the potential to support healthcare and social services by assisting with lifting, mobility, or medication delivery in ways that would augment the capabilities of human caregivers rather than completely replace them.

An important factor driving interest in humanoid robots is that the physical world is designed around the human form factor. If robots are going to be helpful to humans in daily life, their resemblance to humans will benefit their integration into day-to-day activities because they will be better adapted to any given physical human environment than a robot in any other form factor. However, their form factor might also raise unrealistic expectations about their capabilities. As Rodney Brooks, a leading figure in the field of robotics, has noted, "The visual appearance of a robot makes a promise about what it can do and how smart it is. It needs to deliver or slightly over-deliver on that promise or it will not be accepted."[3]

The anthropomorphic design of humanoid robots suggests compatibility with environments built for humans, fueling expectations that they can perform a wide variety of tasks. Recent progress in humanoid robot autonomy has been enabled by advances in data-driven machine learning (ML). The underlying data must be of high quality and is typically obtained from observing the teleoperation of humanoid robots. Improving the quality, volume, and methods for collecting robot teleoperation data will be necessary to achieve human-level precision and dexterity, as well as to improve the autonomy of these robots in more generalizable contexts.

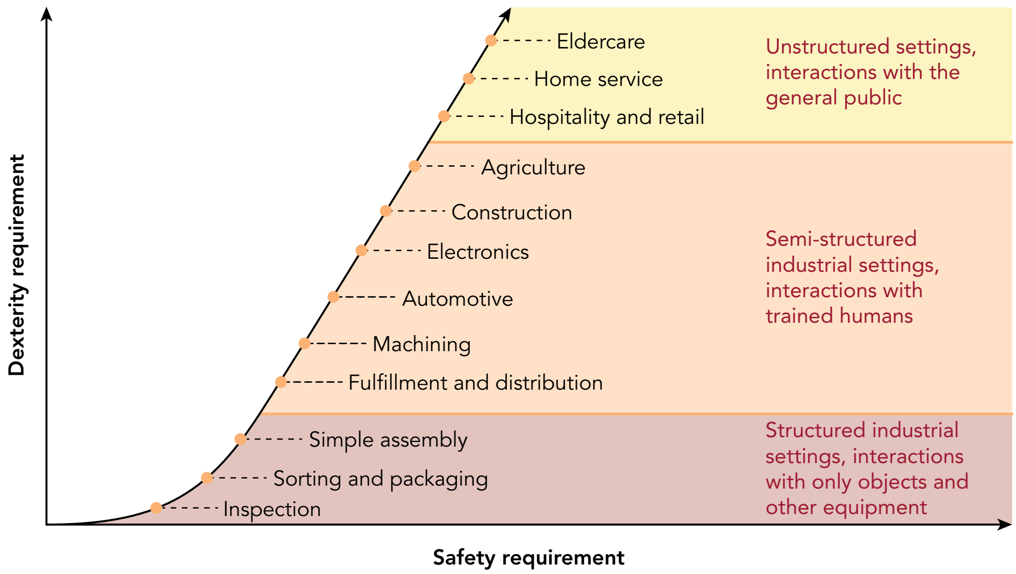
Humanoid robots are not optimally suited for all tasks. Many problems that they can solve—such as material handling or repetitive assembly—can be addressed more effectively with other kinds of robots. In such cases, replicating the full versatility of human movement increases cost and complexity without guaranteeing better performance. With high-value components such as robust actuators, dexterous hands, and force sensors, the high costs of humanoid robots, with average selling prices as high as $200,000 in 2024,[4] have thus far made it difficult to achieve widespread adoption in households. Energy inefficiency, limited battery life, and safety concerns also remain major barriers for everyday use of humanoids in household settings.

While humanoid robots may find specialized roles in industrial and healthcare contexts, wider use of them will depend on multiple factors. Progress in actuation, control, and AI will be critical for making humanoid robots a practical and sustainable solution for real-world applications. As autonomous capabilities improve, it will be particularly important that humanoid robots do not take away from human users' sense of agency, such as their ability to effectively override or alter robots' actions during certain interactions with them.

As shown in figure 8.3, some of today's humanoids are ready for deployment in relatively basic areas such as simple assembly tasks and inspection (the portion of the chart shaded in red), and are nearing levels of dexterity that will allow them to work in the vicinity of trained workers in industrial settings. Humanoid interactions with the general public will require greater dexterity, and in those cases, major technical challenges need to be surmounted to ensure the safety and reliability of humanoids.

Humanoid robots are not optimally suited for all tasks. Many problems that they can solve . . . can be addressed more effectively with other kinds of robots.

FIGURE 8.3 As safety and dexterity improve, humanoid robot deployment will expand to tasks requiring greater human interaction in less structured environments



Source: Adapted from "Humanoid Robots: From the Warehouse to Your House," *Agility Robotics*, blog, July 15, 2025, https://www.agilityrobotics.com/content/humanoid-robots-from-warehouse-to-your-house

Robotics in Warfare

Robotics is significantly reshaping modern warfare, driven by advances in autonomy, communications, and cost-effective robotic technologies. The realities of war in Ukraine and the threat of a conflict in the Taiwan Strait have spurred the Pentagon to reevaluate its combat capabilities. At the core of this reevaluation are autonomous robotic weapons systems. To adapt to this new paradigm, the Pentagon is looking to leverage more intelligent systems and cheap, scalable hardware from start-ups in Silicon Valley and elsewhere.

Examples of types and uses of military robots (which are also known by names such as "uncrewed" or "unmanned" vehicles and drones) include:

- Logistics and last-mile resupply, which can be performed by low-cost unmanned ground vehicles (UGVs) and autonomous convoys that move ammunition, water, and medical supplies under fire, reducing the exposure of human personnel (see figure 8.4). Using robots for casualty evacuation and blood/medicine delivery reduces the risks to medics and pilots.[5]

- Explosive ordnance disposal and route clearance, which can be done by tele-operated and semiautonomous ground robots that can clear mines and improvised explosive devices (IEDs) and inspect hazardous environments, helping to limit personnel's exposure to risk and preserving combat momentum. Robot-based sensing and reconnaissance in areas inaccessible to global positioning systems (GPS) and in subterranean areas (e.g., tunnels and trenches) enables targeting and force protection at lower cost.

- Surveillance drones and robots, which provide real-time situational awareness, persistent monitoring (see figure 8.5), and rapid data collection in contested or dangerous environments while reducing risks to human personnel during reconnaissance, patrol, and targeting operations. Such drones and robots can be deployed on land, at sea, underwater, and in the air.

- Armed drones and robots operated remotely, which enable precision attacks and persistent presence. These unmanned systems—which include UGVs carrying guns or grenade launchers, semiautonomous loitering munitions, seaborne drone boats, and kamikaze quadcopters—extend operational reach, provide rapid and flexible response, and can deliver targeted effects in hostile, contested, or denied environments.

An emerging defense concept integrates many of these robotic platforms with a networked infrastructure that complements the use of legacy crewed systems that are highly capable but also expensive.[6] This approach allows weapons platforms to deploy at scale with reduced human involvement, adapt to evolving combat conditions, and execute tasks such as rapid reconnaissance, precision strikes, and collaborative decision making. These capabilities enhance overall operational adaptability and efficiency and reduce casualties. Lessons from the war in Ukraine, where the high cost of conventional systems constrains their deployment, underscore the value of affordable and quick-to-acquire robotic platforms.[7]

While robots are increasingly being deployed in battlefield environments, the heavy presence of electronic jamming that impedes communication between operators and their remotely controlled platforms poses a challenge. Countering such jamming can be accomplished through creating platforms that are more autonomous (e.g., ones using AI-based capabilities that enable autonomous target recognition or navigation) or by deploying jam-resistant communications (e.g., fiber-optic cables attached to a drone or other platform that unspool as the platform flies through the air and maintain a physical connection to the operator).

Questions also remain around the reliability, accountability, cost, and cybersecurity of highly autonomous systems, and their effectiveness in real-world operations is still being evaluated. Moreover, autonomy in lethal operations introduces unresolved legal and ethical questions, while reliance on interconnected swarms exposes military networks to new vulnerabilities from cyberattacks and electronic warfare.

FIGURE 8.4 THeMIS UGV fifth generation

Source: Milrem Robotics, Wikimedia Commons, CC BY-SA 4.0

FIGURE 8.5 US Army unmanned aircraft system

Source: US Army photo by Pfc. Peter Bannister, Wikimedia Commons

While these and other challenges need to be taken into consideration when thinking about the use of robots by the military, efforts to accelerate the development and strategic integration of robotics in warfare are already helping to redefine combat strategies, increase military effectiveness, and alter geopolitical power dynamics.

Over the Horizon

Future Impacts of Robotic Technologies

MANUFACTURING

The US manufacturing sector, a vital part of the economy contributing over a tenth of US GDP and employing thirteen million people,[8] is facing significant challenges that robotics can help solve. One major issue is a persistent shortage of skilled labor, driven by people retiring and a declining rate of population growth. This scarcity threatens to leave millions of jobs unfilled and could impact the nation's prosperity and security. Widespread adoption of robots in the manufacturing sector is also an important pillar of support for the present push of the US government to emphasize domestic manufacturing by increasing its cost competitiveness compared to foreign manufacturing.

The manufacturing sector's reliance on global supply chains has also made it vulnerable to disruptions, as highlighted during the COVID-19 pandemic. Robotics offers solutions to this challenge through innovations like advanced robotic graspers and collaborative robots, or cobots, which can make manufacturing lines more adaptable and help alleviate labor shortages. These technologies also show promise for manufacturing in extreme environments, like in space or underwater, further enhancing the sector's capabilities and resilience.

THE "NOW" ECONOMY

The "now" economy, which focuses on the near-real-time delivery of goods and services, is leveraging robotics to overcome challenges in logistics and remote service provision. Robots, including delivery drones and autonomous vehicles, are being tested for last-mile deliveries to get products to customers quickly and efficiently. In healthcare, robot-assisted surgeries are becoming more common for remote procedures, while in homes, robots are automating basic services like floor cleaning. In agriculture, autonomous robots are being used as on-demand labor for seasonal tasks such as fruit picking.

However, the widespread adoption of these technologies faces challenges. These include ensuring the safety of delivery drones and privacy issues associated with them, developing robots with adaptable manipulation skills sufficient to handle various types of goods, and creating the necessary networking infrastructure to support reliable remote applications in fields like healthcare.

SUPPORTING AN AGING POPULATION

Robotic technologies are emerging as a crucial solution to the growing challenges of eldercare, which faces a significant shortage of qualified human caregivers. Demand is being driven by a population that is living longer, with a fifth of Americans expected to be over sixty-five by 2030.[9] To address this, robots are being developed to serve as assistive companions that help with daily tasks, exoskeletons that aid mobility, and smaller devices that monitor health and alert professionals to falls or other emergencies. These technologies aim to support human caregivers by handling routine tasks, making eldercare more manageable and accessible.

An aging population also requires more invasive medical care, including surgery. It's estimated that 30 percent of necessary surgeries worldwide go unperformed,[10] and robots can help by automating parts of routine procedures to make them safer and more efficient. Developments in force sensors and haptics (i.e., technology that interacts with human users through touch or other physical sensation) are also enabling telerobotics, which allows doctors to perform surgery on patients from a remote location. This is particularly beneficial for rural and low-income areas where access to specialists is limited.[11]

A major challenge for telesurgery is the complexity of surgical tasks, which varies greatly among patients. The safe implementation of these technologies relies on extensive training and testing, often through simulations, to ensure robots can handle the unpredictable nature of interaction with anatomy.

TACKLING THE HOUSING SHORTAGE

In the face of a housing crisis in the United States marked by high prices and low supply, the construction industry is struggling with a significant labor shortage.[12] Robotic technologies offer a potential solution by increasing productivity and enhancing worker safety. Robots are already commercially available for tasks like bricklaying, framing, and heavy lifting.[13] Beyond housing, robotics can also improve infrastructure projects by automating road paving, inspection, and repair with greater precision. However, integrating these robots into construction sites presents challenges, including the need to ensure they are able to safely navigate unpredictable environments and the necessity of training human workers to collaborate with and maintain these systems.

FOOD PRODUCTION

To keep up with a global population expected to reach ten billion by 2030,[14] food production needs to increase by 50 percent by 2050, a goal complicated by climate change. Robotics offers a solution to streamline food production and processing, with current applications including milking, seeding, and fruit picking. While challenges remain, particularly in tasks requiring high dexterity, like meat carving, the integration of robotics with AI and computer vision is critical. This allows robots to learn complex tasks through reinforcement learning and to capture valuable data on crop health and ripeness. This wealth of information supports precision agriculture—a strategy that uses data to optimize farming practices, reduce fertilizer and water use, and ultimately increase yields and improve soil health, all while helping the industry meet growing global demand.

ADVANCING SUSTAINABILITY

Robotics can play a significant role in advancing sustainable practices across multiple sectors. In renewable energy, robots are essential for the construction and maintenance of solar and wind farms. For example, robots can inspect and repair wind turbines, reducing costs and downtime. They can also help with sustainable resource gathering, such as harvesting lumber with minimal environmental impact and collecting materials from the ocean floor without damaging marine ecosystems. In waste management, AI-powered robots can precisely sort recyclables, making the process more efficient and safer for workers who handle hazardous materials.

Beyond energy and resource management, robotics is also transforming sustainable agriculture and infrastructure maintenance. In farming, technologies like John Deere's See & Spray use ML and cameras to apply herbicides with high precision, drastically reducing chemical use and waste. This technology also helps improve harvesting efficiency and can adapt to the challenges of a changing climate. For infrastructure, robots like Boston Dynamics' Spot can inspect industrial sites and detect gas leaks, preventing failures and improving safety. While the potential for robotics to advance sustainability is clear, continued investment in technology development is crucial to ensure these systems are reliable and can operate safely alongside humans in diverse and unpredictable environments.

Policy Issues

Adoption and Funding

To fully leverage robotics for economic growth and to address labor and supply chain challenges, the United States needs a concerted effort from both the government and the private sector. The US lags behind other leading nations in manufacturing-robot density,[15] with a slower adoption rate, especially among small and midsize businesses (figure 8.6). A key step is to establish clear regulatory guidelines and standards to ensure the safe and effective use of robots in sectors like construction.

Additionally, workforce development can help accelerate the adoption of robots if investments in education and training (e.g., efforts to help workers adapt to working with robots or to transition to new roles) are able to successfully address individual and societal anxieties and concerns about job displacement.

It is also noteworthy that wider adoption of industrial robots can have a significant positive impact on worker safety by moving employees away from dangerous tasks like welding and heavy-duty material handling. Studies indicate a substantial reduction in workplace accidents and fatalities in areas with high robot adoption.[16]

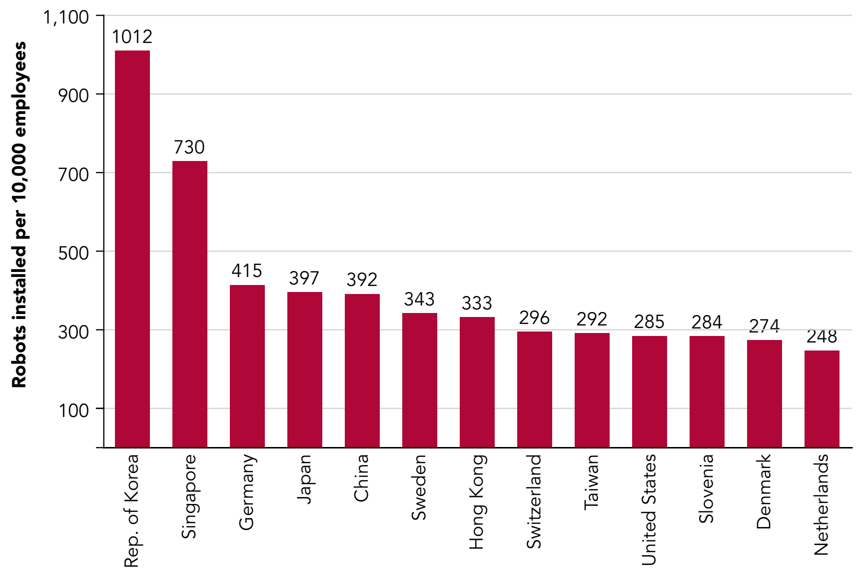
Ultimately, funding of robotics research and development (R&D) remains an issue. Despite efforts such as the National Robotics Initiative (2011–22), more R&D support will be needed if the United States is to make the most of the exciting and transformative opportunities that robotics offers.

Privacy and Consent

The exponential growth of data collection by robots in homes and hospitals for both training and operational purposes raises significant concerns about how this personal and sensitive information is handled and secured. Just as health information is heavily regulated to protect patient privacy, policies must be developed to safeguard the vast amounts of data that will be used in the future to train and operate robotic systems.

For example, standards for data privacy will need to be put in place if humanoid robots are to operate unsupervised around vulnerable individuals in

FIGURE 8.6 The United States lags behind many other countries in manufacturing-robot adoption



Source: Adapted from "Global Robotics Race: Korea, Singapore and Germany in the Lead," news release, International Federation of Robotics, January 10, 2024, https://ifr.org/ifr-press-releases/news/global-robotics-race-korea-singapore-and-germany-in-the-lead homes or care facilities, as there are currently no defined regulations or standardized certification processes for this. The use of data or AI models developed in countries like China may have reduced utility or fitness for purpose when applied within the US environment because of regulatory and contextual mismatches (e.g., different demographics in the populations that supply the data that are collected) and differing privacy regulations and data collection standards.

Inclusion and Integrity

The potential for bias in datasets used to train robots could lead to serious, harmful outcomes. For instance, a surgical robot might be less effective operating on patients from one demographic group if its training data is predominantly from another demographic group. To prevent such dangerous scenarios, it is critical to promote and enforce standards that ensure robot-training datasets accurately reflect relevant characteristics of different groups in the population at large.

Safety

The safety of robots, including physical and cybersecurity aspects, is a critical legal and ethical issue. A key challenge for public acceptance is whether safety standards should mirror human performance (or even above-average performance) or approach near perfection. The former is easier to achieve from a technical perspective, but public acceptance of such a standard is uncertain. Responsible adoption requires clear performance standards and robust cybersecurity, particularly in sensitive sectors like healthcare and national security.

Internationally, ISO 10218, a major update to the global standard for industrial robot safety, was released in early 2025.[17] This update provides clearer guidelines for robot manufacturers and integrators, especially concerning collaborative robots and cybersecurity, with forthcoming US versions of the standard expected to align with these changes.

Supply Chain

The robotics supply chain is central to robotics advancement, and the United States' access to key inputs is vulnerable to disruption due to concentrated dependencies in foreign countries and weak domestic capacity. For example, China dominates mining and processing of rare earth elements used in the high-strength permanent magnets needed for robotic motors and actuators.

Most of the key robot components described at the beginning of this chapter are produced at scale in China—and even when they are designed elsewhere, manufacturing and assembly often happen in Chinese factories due to lower costs and established infrastructure. Many robotics companies (including those from the United States, Europe, and Japan) source parts or assemble products in China because their suppliers are already there. Once a supply chain is concentrated, moving it is both costly and disruptive.

The robotics supply chain is central to robotics advancement, and the United States' access to key inputs is vulnerable to disruption.

Policy Activities to Promote Robotics

Policy in the United States regarding robotics in the past year has been heavily influenced by a new focus on AI and its role in national competitiveness, particularly against the backdrop of a new presidential administration. For example, as noted in chapter 1, on artificial intelligence, the Trump administration released America's AI Action Plan in July 2025.[18] The plan focuses on accelerating AI innovation, building domestic computing infrastructure, and leading in international diplomacy and security issues related to AI. However, the action plan includes little mention of robotics, relegating the topic to a short section on manufacturing.

There is a significant push from both the government and the robotics industry to establish a comprehensive national robotics strategy. Industry leaders and organizations like the Association for Advancing Automation (A3) have called for a dedicated federal office, tax incentives, and expanded workforce training programs to ensure the United States remains a leader in robotics. This comes as a response to global competition, particularly from countries like China, which has its own national strategy to lead in high-tech manufacturing, called "Made in China 2025."

NOTES

1. Ralph Lässig, Markus Lorenz, Emmanuel Sissimatos, Ina Wicker, and Tilman Buchner, "Robotics Outlook 2030: How Intelligence and Mobility Will Shape the Future," Boston Consulting Group, June 28, 2021, https://www.bcg.com/publications/2021/how-intelligence-and-mobility-will-shape-the-future-of-the-robotics-industry.

2. Some examples of soft robots can be found at Alberto Paitoni Faustinoni, "Soft Robotics: Examples, Research and Applications," *Robotics* 24 (February 2023), https://robotics24.net/blog/soft-robotics-examples-research-and-applications/.

3. Rodney Brooks (professor emeritus, Massachusetts Institute of Technology), in his plenary talk at the 2025 Stanford Human-Centered AI Institute Spring Conference on Robotics.

4. Morgan Stanley Research, "Humanoids: A $5 Trillion Global Market," April 29, 2025, https://www.morganstanley.com/articles/humanoids-5-trillion-global-market.

5. "Ukraine Soldiers Use Ground Robots Like Adaptable Lego Sets," *Business Insider*, July 5, 2025, https://www.businessinsider.com/ukraine-soldiers-use-ground-robots-like-adaptable-lego-sets-operator-2025-7.

6. For networks, see https://dodcio.defense.gov/Portals/0/Documents/DoD-C3-Strategy.pdf; for robotic platforms, see https://www.scsp.ai/wp-content/uploads/2024/12/DPS-Joint-Warfighting-Concept-2034-44-.pdf.

7. Andrey Liscovich, cofounder of Ukrainian Defense Fund, personal communication with Luke Hyman, June 16, 2025.

8. "Manufacturing in the United States," National Association of Manufacturers, accessed September 26, 2025, https://nam.org/mfgdata/.

9. America Counts Staff, "By 2030, All Baby Boomers Will Be Age 65 or Older," US Census Bureau, December 10, 2019, https://www.census.gov/library/stories/2019/12/by-2030-all-baby-boomers-will-be-age-65-or-older.html.

10. One study estimates that about 143 million additional surgical procedures are needed in low- and middle-income countries each year to save lives and prevent disability, while 313 million such procedures are performed each year. See John G. Meara, Andrew J. M. Leather, Lars Hagander, et al., "Global Surgery 2030: Evidence and Solutions for Achieving Health, Welfare, and Economic Development," *The Lancet* 386, no. 9993 (2015): 569–624, https://doi.org/10.1016/S0140-6736(15)60160-X.

11. Aashna Mehta, Jyi Cheng Ng, Wireko Andrew Awuah, et al., "Embracing Robot Surgery in Low- and Middle-Income Countries: Potential Benefits, Challenges, and Scope in the Future," *Annals of Medicine and Surgery* 84 (December 2022): 104803, https://doi.org/10.1016/j.amsu.2022.104803.

12. Troup Howard, Mengqi Wang, and Dayin Zhang, "How Do Labor Shortages Affect Residential Construction and Housing Affordability?," Haas School of Business, University of California–Berkeley, April 2023, https://www.haas.berkeley.edu/wp-content/uploads/Howard_Wang_Zhang_Housing-Supply-and-Construction-Labor-Dayin-Zhang.pdf.

13. "SAM," Construction Robotics, accessed September 16, 2024, https://www.construction-robotics.com/sam-2/.

14. "Global Issues: Population," United Nations, accessed September 16, 2024, https://www.un.org/en/global-issues/population#.

15. "US Ranks 10th in Robot Density," *Assembly*, January 17, 2024, https://www.assemblymag.com/articles/98270-us-ranks-10th-in-robot-density.

16. Marco De Simone, Dario Guarascio, and Jelena Reljic, "The Impact of Robots on Workplace Injuries and Deaths: Empirical Evidence from Europe," *SSRN*, February 13, 2025, https://doi.org/10.2139/ssrn.5136996.

17. ISO 10218-1:2025 (Safety requirements for industrial robots) and ISO 10218-2:2025 (Safety requirements for industrial robot applications and robot cells) are available at https://www.iso.org/standard/73933.html and https://www.iso.org/standard/73934.html, respectively.

18. Winning the Race: America's AI Action Plan, Executive Office of the President, July 2025, https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.

STANFORD EXPERT CONTRIBUTORS

Dr. Allison Okamura
SETR Faculty Council, Richard W. Weiland Professor in the School of Engineering, and Professor of Mechanical Engineering and, by courtesy, of Computer Science

Dr. Mark Cutkosky
Fletcher Jones Professor in the School of Engineering and Professor of Mechanical Engineering

Dr. Mac Schwager
Associate Professor of Aeronautics and Astronautics and, by courtesy, of Computer Science

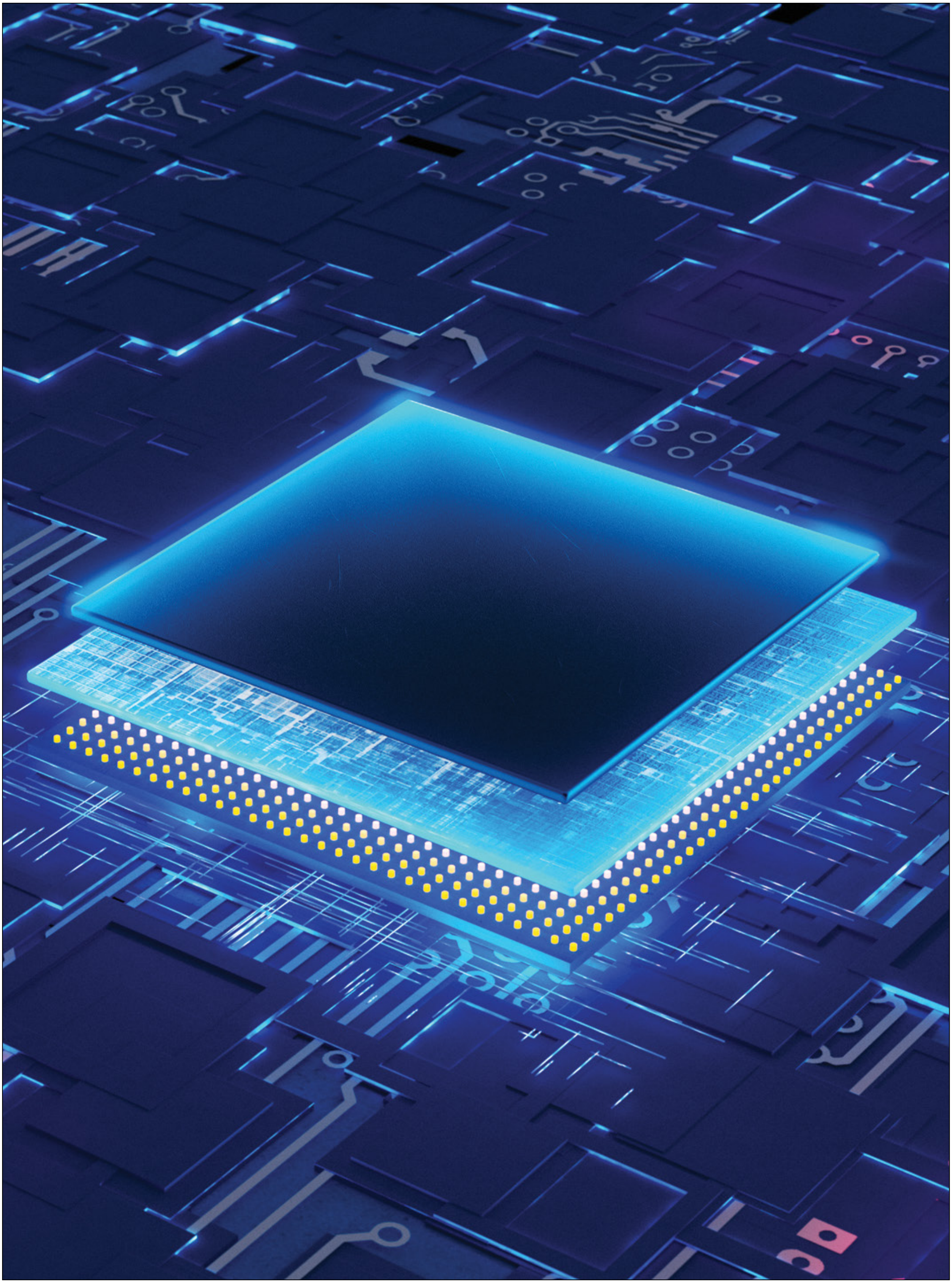
Dr. Renee Zhao
Assistant Professor of Mechanical Engineering and, by courtesy, of Materials Science and Engineering

Dr. Monroe Kennedy
Assistant Professor of Mechanical Engineering and, by courtesy, of Computer Science

Dr. Cosima du Pasquier
SETR Fellow and Postdoctoral Scholar in Mechanical Engineering

Brian Vuong
SETR Fellow and PhD Student in Mechanical Engineering

Luke Hyman
SETR Fellow and PhD Student in Mechanical Engineering



09

SEMICONDUCTORS

KEY TAKEAWAYS

- The growing demand for artificial intelligence (AI) and machine learning is driving innovations in chip fabrication, along with advances in memory technologies and high-bandwidth interconnects such as photonic links, all of which are essential for enhancing computational power, managing energy efficiency, and meeting the increasing data needs of modern applications.

- Semiconductor manufacturing is the most precise manufacturing process that exists. It is used to advance work in energy and biotechnology in addition to information technology and AI.

- Strategic technology containment efforts directed against China help constrain Chinese capabilities in the short term. However, they are likely to drive China into a technology posture that is considerably more decoupled from the West and hence less vulnerable to Western pressure in the future.

Overview

Semiconductors, often in the form of microchips, are crucial components used in everyday physical devices, from smartphones and toasters to cars and lawn mowers. Chips control heating and cooling systems, elevators, and fire alarms in modern buildings. Traffic lights are controlled by chips. On farms, tractors and irrigation systems are controlled by chips. Modern militaries could not function without chips in their weapons, navigation devices, and cockpit life-support systems in fighter jets. The list goes on and on—in every aspect of modern life, chips are essential.

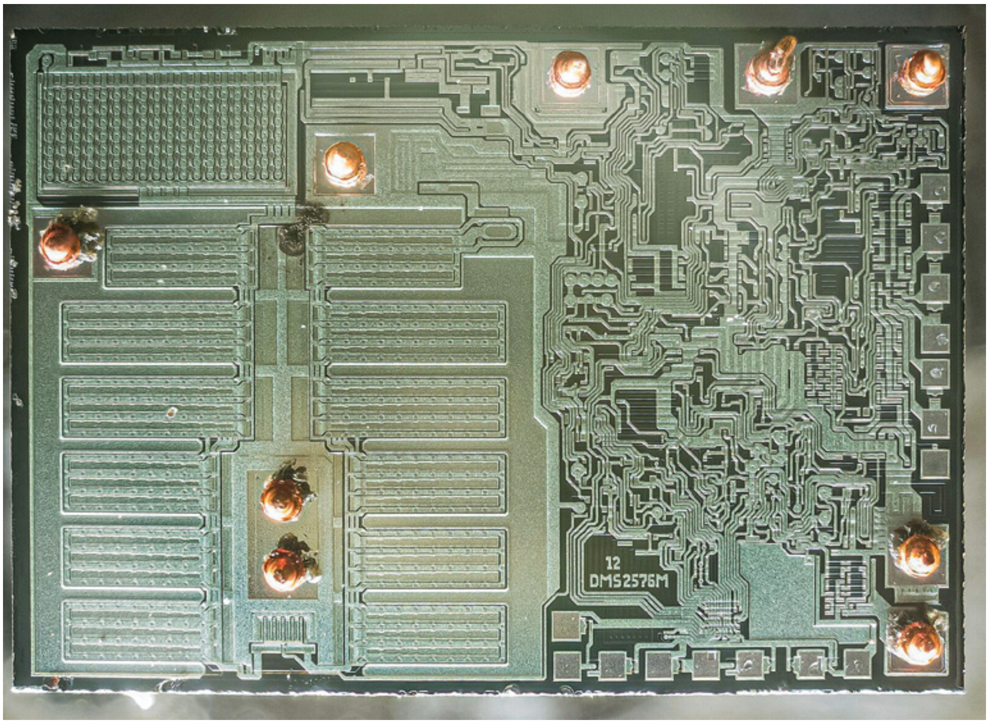
Most chips are involved in the handling of information. Different types of them are specialized for different tasks. Some are processor chips that ingest data, perform computations on the data, and output the results of those computations. Memory chips store information and are used with processors. Still other chips act as interfaces between digital computations and the physical world. In all these cases, some amount of energy is needed to represent each bit of information inside a chip. The magic of chips is that it takes several orders of magnitude less energy to represent information inside one than it takes to do so outside it (e.g., in wires leading to and from the chip). This means that, in a multi-chip system, much more energy and chip space are required for data moving between chips than for data that remains on a chip; this is one of the driving forces to integrate more functions on a single chip.

As chip fabrication technologies improve, it takes less energy and chip space to represent a given bit of information; hence, processing those bits becomes more energy efficient. This phenomenon is what has enabled the semiconductor industry to pack more processing power on chips over time—it enables designers to create chips that do more complex processing (see figure 9.1). However, the cost of designing them also increases with their complexity.

Recently, however, the energy costs associated with the hardware that holds information on a chip have been falling more slowly, and the cost of manufacturing per unit area has increased. This means that the cost and energy advantages of scaling have nearly stopped. As a result, researchers have been investigating other ways to improve computer technology and to deal with the problem of high design costs.

Since the best technologies for performing different chip functions are themselves different, systems still need to use different chips for those functions. Finding new ways to manage the inefficiency of information movement in and among chips, along with the issue of high design costs, is a central focus of

FIGURE 9.1 Improving fabrication has enabled the creation of more complex chips



Source: Wikimedia Commons, CC BY-SA 4.0

FIGURE 9.2 Chip fabrication requires large factories that can produce chips at scale

Source: IM Imagery/Shutterstock.com

research on semiconductors. Further improvement will take the form of innovation in design, materials, and integration methods.

Two aspects of chips are important for the purposes of this report. They must be designed and then fabricated (i.e., manufactured), and each function calls for different skill sets. Chip design is primarily an intellectual task that requires tools and teams able to create and test systems containing billions of components. Fabrication is primarily a physical effort that requires large factories, or fab facilities, that can produce chips by the millions and billions—and can cost billions of dollars and take several years to build from the ground up (see figure 9.2).

Fabrication integrates many complex processes to produce chips. Each one requires substantial expertise to master and operate, and the integration of all of them requires still further expertise. For these reasons, modern fabrication plants are operated by workforces with a substantial number of people trained in engineering.

Fabrication also entails a significant degree of process engineering to continue to improve process technology and to achieve stringent manufacturing standards. For example, the "clean rooms" in which chips are made require air that is a thousand times more particle-free than the air in a hospital operating room.[1]

Because chip design and chip fabrication are so different in character, only a few companies, such as Intel, do both. However, Intel is in trouble, and some technology analysts and former Intel directors think it should split its design and fabrication groups.[2] Many businesses specialize in design, including Qualcomm, Broadcom, Apple, and Nvidia. Such companies are called "fabless" in recognition of the fact that they do the design work and outsource fabrication to others—a strategy based on the theory that the former activity has higher profit margins than the latter.

Today, the "others" being outsourced to usually refer to one company: Taiwan Semiconductor

Manufacturing Company (TSMC), by far the world's largest contract chip-manufacturing company. In 2024, TSMC controlled over 60 percent of the world's contract semiconductor manufacturing and 90 percent of the world's advanced chip manufacturing.[3] Samsung, in South Korea, is a distant second, with around 13 percent of the world's chip manufacturing.[4] United Microelectronics Corporation, also based in Taiwan, ranks third at about 6 percent.[5]

By contrast, US chip-manufacturing capacity has lost significant ground. Fabrication plants in America accounted for 37 percent of global production in 1990, but their share dropped to just 12 percent by 2021.[6] Industry concentration, low US production capacity, and geopolitical concerns about China's intentions toward Taiwan mean the global supply chain for chips will remain fragile for the foreseeable future, despite the passage of the Creating Helpful Incentives to Produce Semiconductors (CHIPS) and Science Act of 2022 (discussed in more detail later in this chapter).

The strength of semiconductor manufacturing affects more than just information technology. It is also the most precise manufacturing method on the planet and is now driving innovations in areas ranging from neuroscience and synthetic biology to energy and lighting. While many of these applications don't need the most advanced processing technology, they do require access to semiconductor fabrication and fabrication expertise.

Key Developments

Moore's Law, Past and Future

For over half a century, information technology has been driven by improvements in the chip fabrication process. In 1965, Intel cofounder Gordon Moore observed that the cost of fabricating a transistor was dropping exponentially with time—an observation that has come to be known as Moore's law. It's not a law of physics but rather a statement about the optimal rate at which economic value can be extracted from improvements in the chip fabrication process.

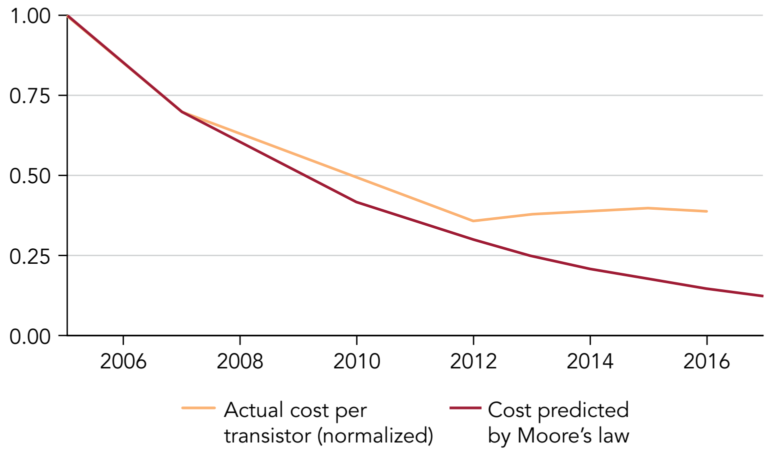
Although Moore's law is often stated as the number of transistors on a chip doubling every few years, historically what drove this scaling was that the cost of making a chip was mostly independent of the number of components on it. This has meant that every few years, a chip whose size and cost remains approximately the same will have twice the number of transistors on it.

Moore's law scaling (i.e., the exponential increasing of the number of transistors on a chip) meant that each year one could build last year's devices for less money than before or could build a more powerful system for the same cost. This scaling has been so consistent that it is widely believed that the cost of computing will always decrease with time. It's an expectation so pervasive that in almost all fields of work, people are developing more complex algorithms to achieve better results while relying on Moore's law to rescue them from the consequences of that additional complexity.

But the future will not look like the past. As the complexity of chips increases, the traditional benefits associated with Moore's law scaling are diminishing, leading to rising costs in chip manufacturing. As figure 9.3 shows, the actual cost of a chip per transistor (represented by the solid orange line) was tracking the cost predicted by Moore's law (the solid red line) relatively well from 2004 to 2012.[7] However, the actual cost per transistor started to level off around 2012, and it has not kept up with Moore's law predictions since then.[8]

Historically, advances in technology have come from shrinking the size of the transistors and the wires that connected them, and the name of the technology was derived from the smallest feature in the design (e.g., a 130-nanometer [nm] chip was one in which the smallest feature was 130 nm across).

FIGURE 9.3 Cost per transistor over time



Source: Adapted from Steve Mollenkopf, "Our Future Is Mobile: Accelerating Innovation After Moore's Law," presentation at the Electronics Resurgence Initiative Summit, Detroit, MI, July 15–17, 2019, https://eri-summit.darpa.mil/docs/Mollenkopf_Steve_Qualcomm_Final.pdf

However, over a decade ago, it became increasingly difficult to boost circuit density (as measured by the number of circuits per square centimeter). Other approaches had to be found to do so, including having transistors use the vertical dimension to decrease their area.

Using the vertical dimension, along with other, more complex processing techniques (such as the use of new materials), enabled circuit density to continue to increase. But technology marketers continued to use a shrinking distance to characterize newer generations of more densely packed circuits, despite that distance no longer representing anything physical. In other words, the name became a marketing device (or, to put it more kindly, a generational technology label), even though it still sounded like it referred to a distance.

When the label had actual physical meaning, such as indicating the size of a feature, that number (e.g., 130 nm) could be used to make inferences about chip performance, such as the actual cost per compute or the energy needed for a computation. But once it became a marketing term, the connection between the label and the chip's performance was broken.

Against this backdrop, the past year has witnessed significant advancements and challenges in the semiconductor industry. For example, increasing demand for computing power driven by artificial intelligence (AI) and machine learning (ML) applications (as discussed in chapter 1, on artificial intelligence) has led to a surge in the development of, and demand for, advanced graphical processing units (GPUs). This has created a strain on both production and energy resources.

Traditional processors are not the best approach for the intensive computational tasks required by modern AI algorithms. As a result, there has been a significant investment in developing GPUs and other specialized processors designed specifically for AI workloads. This shift is reshaping the semiconductor industry, emphasizing the need for high-performance, energy-efficient computing solutions.

The increased density, provided by scaling and advanced packaging together, has enabled advanced GPU systems to connect a massive amount of compute in a small space, increasing performance. (Advanced packaging refers to the combination of circuitry from multiple semiconductor chips or dies into a single, compact electronic package, as in 2.5-dimensional [2.5-D] integration.) One example of this is Nvidia's GB200 NVL72 system. Unfortunately, the industry has been unable to shrink power and fabrication cost as rapidly as in the past, and thus these machines are both extremely expensive and very power hungry. They dissipate ten times the power of systems of a few years ago. They also require innovations in order to get the power into a system and to get the heat out. This surge in demand for high performance underscores the importance of finding innovative solutions to meet computational needs in the most energy-efficient and simple (reducing power and cost, respectively) way possible.

Addressing this challenge requires specialized hardware that is extremely efficient in computing the results needed by today's AI applications. This is the only known way to grow computation performance per dollar and per watt. Today, all computing devices used for AI applications, including GPUs, contain this type of specialized hardware. Nvidia reports its optimizations reduced the required energy to perform computations by a thousand times.[9] But even with these optimizations, current computing systems are dissipating a large amount of heat (over 100 kilowatts [kW] per rack). For comparison, a typical US household uses electric power at the rate of about 1 kW averaged throughout the year—and the power used in future systems is projected to keep increasing.

Chiplets and 2.5-Dimensional Integration

Integrating an enormous amount of compute and memory on a single piece of silicon is desirable for maximum energy efficiency. However, this poses two problems. First, some of today's most demanding applications require more computing resources than can be manufactured on a single piece of silicon. Second, the manufacturing processes for processors and memory are very different, so these two components can't be placed on the same piece of silicon.

One solution to these challenges is the use of chiplets and 2.5-D integration. Here, rather than forcing everything onto a single piece of silicon, multiple silicon pieces are connected together on an interposer (defined in more detail below) to create a much larger superchip.

This superchip combines processors and memory using chiplets and 2.5-D integration, which leverage different manufacturing technologies to optimize each component. Chiplets—functional blocks of silicon—can be combined in various ways, enabling vendors to tailor systems to customer needs. Central to 2.5-D integration is the interposer, a specialized substrate that connects chiplets and facilitates faster, more energy-efficient communication than traditional circuit board wiring. By allowing high-density memory, high-performance compute units, and communication chips to reside side by side, this approach boosts bandwidth, performance, and power efficiency while reducing the need for full integration on a single chip.

These superchips can contain both memory and processors. Representing a significant shift from traditional monolithic chip design, 2.5-D integration increases per-transistor cost because both the chips that the transistors are on and the substrate must be manufactured. Nevertheless, 2.5-D integration enables semiconductor companies to create a set of building-block chiplets that can be combined in various ways, allowing for the range of products with different performance characteristics mentioned earlier. This strategy allows companies to better tailor their products to specific application domains and more effectively monetize their silicon investments, ultimately leading to increased product diversity and market responsiveness.

Given the changing economics of scaling, the use of chiplets reduces costs overall and allows for more customized solutions. Exemplifying this strategy is the approach taken by semiconductor company AMD, which involves keeping components that transfer data to and from devices in older technology nodes while advancing core computing resources with the latest processes. Moreover, this modular strategy facilitates the integration of emerging technologies such as photonics (discussed in greater detail later in this chapter), which can significantly enhance communication speeds and bandwidth within and between chips.

High-Power Density

Moving the compute and memory elements closer together improves system performance. But all such systems generate heat—and combining more elements on a single chip increases the amount of heat that must be shed during system operation.

For example, in its NVL72 system, Nvidia packs seventy-two B200 2.5-D mega GPUs into a rack and uses its high-performance NVLinks technology to connect all the GPUs to each other, thus forming a supercomputer pod with immense computational power and bandwidth. Each GPU in these setups is approximately ten times more powerful than a consumer one, dissipating a kilowatt of power each. Two of these are placed on a board that dissipates 2.7 kW, and the boards are placed in a rack that dissipates a total of 120 kW. Four to eight of these racks can be connected together with longer range links to create a super pod that dissipates 0.5 to 1 megawatt of heat. (For comparison, a 2,500-square-foot house might require a furnace that generates about 30 kW of heat.)

Thus, thermal management stands out as a critical issue. Historically, computing equipment was cooled by blowing cold air through a machine. But moving cold air is insufficient to remove this level of heat, and high-performance, compute-intensive machines must use liquid flowing through cooling plates to deal with it. The heat absorbed by the liquid must then be dissipated somewhere else, usually into the air using large air-conditioning units on the roof of a building. Effective thermal-management solutions, such as advanced cooling techniques and materials with high thermal conductivity, are essential to maintaining performance and reliability in high-performance computing systems.

Need for High Bandwidth

As the prior example showed, while 2.5-D integration helps provide local bandwidth, modern systems are large enough that they require many of these highly integrated superchips. Communicating information between these systems is therefore critical. AI-training models must handle vast amounts of data, and high-speed interconnects, such as those employed in Nvidia's B200 systems, play a critical role in facilitating the rapid transfer of data between compute units and memory.

Traditionally, this communication between the chips in a rack is done on electrical wires embedded in the boards that the chips connect to. As the communication rates have continued to rise, the physical limitations of traditional electrical interconnects are one of the primary barriers to improving bandwidth.

To overcome this bandwidth limit in communicating information from a chip on one board to a chip on another board, researchers are developing "flying cable" connectors. These connectors are placed directly on the top side of a superchip and enable a high-performance cable to be attached directly to the chip, while the other end is connected directly to the connector to the other board. This cable is built to have the best possible signal transmission properties. Researchers are experimenting with both electrical and optical cables, providing an ability to increase interface speeds above the 100 billion bit per second per wire rate used today.

Memory Technology Developments

Memory technology continues to evolve, with innovations in both stacking and new materials. Techniques such as stacking multiple layers of flash memory (i.e., memory that retains its contents even when power is shut off) are pushing the boundaries of what is possible, enabling higher density and better performance. These advancements are crucial for supporting the growing data needs of modern applications, from AI to big data analytics.

Dynamic random-access memory (DRAM) and flash memory technologies have both seen significant advancements in recent years, but the associated increase in manufacturing cost means there has been only modest improvement in cost per bit. The development of three-dimensional (3-D) structures (e.g., vertical DRAM transistors) has allowed for continued scaling of memory density by overcoming the physical limitations of traditional planar transistors. Three-dimensional packaging has enabled the production of memory devices with higher capacity and improved performance, known as high-bandwidth memory (HBM).

Dynamic random-access memory (DRAM) and flash memory technologies have both seen significant advancements in recent years, but the associated increase in manufacturing cost means there has been only modest improvement in cost per bit. Both DRAM and flash memory moved to 3-D structures decades ago and have had to use increasingly complex structures to scale the memory cell size. Chip stacking has also been used for many years to increase the number of memory bits shipped in a single package. What has changed recently is the growth of the HBM market. These memories require a more complex chip-stacking technology called through-silicon vias (TSV), which needs many wires to run vertically through the chips.

As memory technologies scale, maintaining performance and reliability becomes increasingly challenging. For DRAM, issues such as leakage currents and quantum effects limit the scalability of capacitors and transistors. To address these challenges, researchers have developed advanced manufacturing techniques for the creation of complex 3-D structures that enable increased storage density while maintaining the required electrical characteristics.

The boom in AI computing has also affected the DRAM industry. The enormous compute power of today's specialized ML systems means enormous amounts of data per second, or data bandwidth, are required to keep them busy.[10] This need for high-bandwidth memory has created the growing market for HBM mentioned earlier. As a result, the South Korean firm Hynix, initially the only manufacturer of HBM, has grown into the largest DRAM manufacturer, overtaking Samsung, which led the market for many years.[11] China has also invested heavily in DRAM and flash memory production, with Chinese companies selling less advanced parts at very low (possibly below-cost) prices, forcing Samsung and Hynix out of that business.[12] Recently, these Chinese companies have been directed to move to the more advanced memory devices.[13]

Similarly, NAND flash memory—the most common type of flash memory—transitioned to a 3-D cell design in the mid 2010s and has been scaling density by scaling the number of layers in 3-D transistor stacks for the past decade. However, this approach requires sophisticated manufacturing processes to ensure the reliability and performance of the resulting memory devices.

Emerging memory technologies, such as magnetoresistive random-access memory[14] and phase-change memory,[15] are also gaining traction as an alternative to today's embedded nonvolatile technology. These technologies offer advantages in terms of speed, endurance, and energy efficiency, making them attractive alternatives to traditional embedded nonvolatile memory solutions. (Nonvolatile memory retains its contents even when power is turned off.)

Further innovations in memory technology are critical for enabling the continued growth of data-intensive applications. From AI-training models to cloud computing and big data analytics, modern applications require vast amounts of memory to store and process data efficiently.

Over the Horizon

The semiconductor industry is poised for significant advancements in coming years, driven by the growing demands of AI, especially ML, and high-performance computing. The introduction of new technologies, such as 2.5-D integration, chiplets, and photonic interconnects, is expected to play a crucial role in meeting these demands. These innovations will help to enhance performance, increase bandwidth, and improve energy efficiency, addressing the limitations of traditional semiconductor designs.

Emerging memory technologies and advanced manufacturing techniques are also critical for the industry's growth. Innovations in memory stacking and integration with processors will improve data-transfer speeds and reduce latency, meeting the increasing data requirements of modern applications. The development of advanced materials and transistor architectures will further push the boundaries of semiconductor capabilities, enabling continued miniaturization and enhanced performance.

Three-Dimensional Heterogeneous Integration

As noted above, advanced chip designs sometimes use 3-D structures. Today, these designs are limited to a variety of niche applications, such as HBM and high-performance computing. These 3-D structures are the result of a fabrication technique known as 3-D heterogeneous integration. This is different from 2.5-D integration, where different chiplets are placed on a common substrate. Rather, true 3-D heterogeneous integration is a semiconductor manufacturing technique that involves the vertical stacking of different electronic components, such as processors and memory, with vertical interconnect between them. The heterogeneous aspect means that these stacked components can be made from different materials and technologies optimized for their specific functions.

For example, a processor made with one type of fabrication can be stacked with memory made from another, with each using the most suitable technology for its purpose. This approach has the potential to improve performance and efficiency by reducing the distance data travels between components, making devices faster and more compact—albeit at the cost of a more complex fabrication process and a harder heat-dissipation task for the resulting chips.

A variety of challenges need to be overcome for 3-D heterogenous integration to become more widely used. These include thermal management, mechanical stress and reliability, manufacturing complexity and cost, interconnect reliability, and design complexity. Many of these issues are also present with traditional 2-D and 2.5-D integration, but vertical stacking creates new failure modes that do not exist or are much less severe in traditional 2-D chips.

Photonic Links and Components

The distance that a high-performance electrical data-transmission link can span has been shrinking as its data-bandwidth has increased. Photonic (light) links are now used for longer-distance communications. Photonics is the optical analog of electronics—the latter use electrons for signaling and carrying information, while photonics use photons (light) for the same purposes. Innovations such as silicon photonics are emerging, making photonic links attractive for much shorter distances, including some chip-to-chip communications.

Silicon photonic links have the potential to reduce energy consumption and increase bandwidth in data centers and long-distance data transmissions that are not already photonic.[16] Furthermore, they can handle different wavelengths simultaneously on a single optical fiber. This enhances data-carrying capacity and makes photonics an attractive solution for high-performance computing and data center applications. By replacing electrical interconnects with optical ones, data centers can reduce the amount of energy required for transmitting data, leading to lower operational costs and a smaller environmental footprint. Such advantages of photonics have always been drivers of research in this area, but the recent rise in the demand for power-hungry AI-enabled applications has created even more impetus for such research.

Integrating photonic components with silicon-based technologies is challenging as a result of material incompatibilities; for example, efficient light-emitting materials like III–V semiconductors do not integrate well with silicon. (A III–V semiconductor is made by combining boron, aluminum, gallium, or indium with nitrogen, phosphorus, arsenic, or antimony.) While useful for light detectors, silicon is ineffective for light emission, complicating the scalable integration of these technologies at the chip or circuit board level. Overcoming these challenges is crucial for realizing the full potential of photonic links in large-scale, low-energy applications.

Applications-Specific Optimization

Finally, as Moore's law reaches its limits, future improvements in computing will rely more on optimizing algorithms, hardware, and technologies for specific applications rather than on general technology scaling. This requires innovation across the entire technology stack, from materials to design methods. However, the industry faces a paradox: The need for radical innovation conflicts with the high costs and long timelines of chip development, which can exceed $100 million and take over two years.

To address this, the industry must make system-design exploration easier, cheaper, and faster. Researchers are working to ensure that specific design changes to a chip do not require redesign of the entire chip. Solutions include enabling software designers to test custom accelerators without deep hardware knowledge and developing tools for application developers to make small hardware extensions to base platforms. This approach, described in more detail in the inaugural *Stanford Emerging Technology Review* (*SETR* 2023), depends on the involvement of major technology firms, who would need to participate in an app store–like model for hardware customization, balancing open innovation with profit motives.

Policy Issues

Talent

A critical challenge facing the US semiconductor industry is its significant talent shortage, particularly in hardware design and manufacturing. For example, the Semiconductor Industry Association projects the number of jobs in the sector in the United States will grow by nearly 115,000 by 2030, to total approximately 460,000.[17] Moreover, it estimates that roughly 67,000, or 58 percent, of these new jobs risk going unfilled at current degree-completion rates. Looking at just the new jobs that are technical in nature, the percentage at risk of going unfilled is higher, at 80 percent. Almost two-thirds of the unfilled jobs would require at least a bachelor's degree in engineering.[18]

The pipeline of college graduates interested in semiconductors is also troubling. While student interest in hardware seems to be increasing, recent actions, including the voiding of up to $7.4 billion in CHIPS Act funding[19] and the cutting of government funding for research in general, will inevitably shrink the number of new graduates in this area.

Since appropriately trained people are the only real source of new ideas, this trend does not bode well for the industry. Addressing this issue requires more and even closer collaboration among educational institutions, industry, and government to develop programs that attract and train the next generation of semiconductor engineers and researchers.

Strategic Technology Containment

The primary objective of actions taken under this rubric is to restrict China's access to high technology to preserve Western advantages in innovation, industry, and defense. For example, the United States has intensified export controls and revocation of export authorizations targeting Chinese semiconductor firms and key software and equipment for design and fabrication. These efforts aim to restrict China's ability to develop the most advanced semiconductors (currently 5 nm and 3 nm technologies). Restricted technologies include top-tier lithography machines, high-performance computing chips, and electronic design automation software.[20]

In December 2024, the US Department of Commerce released a set of export-control rules to further impair China's ability to produce advanced semiconductors, specifically targeting, among other things, certain semiconductor manufacturing equipment.[21] This equipment included lithography tools using extreme ultraviolet light (EUV), thus effectively blocking deliveries of ASML Holding's most advanced EUV systems to China.[22] ASML is a Dutch multinational corporation that develops and manufactures advanced photolithography machines used in semiconductor fabrication and the only company worldwide that produces EUV lithography systems for this purpose.

In addition, the US-led Clean Network program, launched in 2020 in the first Trump administration, sought to prevent Chinese telecommunications carriers and suppliers from accessing or influencing sensitive US telecommunications infrastructure and data networks.[23] Further, the Federal Communications Commission banned the US sale of certain communications equipment from Chinese technology firms such as Huawei and ZTE.[24]

Such actions have had a disruptive impact on China's semiconductor ecosystem,[25] at least in the short term. The inability of Chinese firms to access key tools for next-generation chip production has triggered supply chain delays, steep price increases, and direct setbacks, resulting in workforce reductions and postponed factory expansions. China's progress in producing state-of-the-art chips has been impeded, forcing it toward legacy technologies; its broader ambitions in AI and advanced computing have also been impacted.

In response, China has launched a comprehensive effort to achieve semiconductor self-sufficiency. This includes massive subsidies, accelerated investment, and reforms supporting indigenous chip design and manufacturing innovation. It has further supported research in the field (e.g., producing twice as many chip design papers than the United States) and made advances in materials like 2-D transistors and carbon nanotube chips. Finally, it has undertaken a variety of efforts to circumvent Western containment measures, including the use of smuggling and shell companies to purchase equipment and chips.

Thus, while Western technology containment efforts have effectively slowed immediate Chinese advances, they may have the unintended impact of decreasing Chinese dependence on Western technology.

To further discourage technology containment efforts, China also retains the option to retaliate against

A critical challenge facing the US semiconductor industry is its significant talent shortage, particularly in hardware design and manufacturing.

nations pursuing them. For example, in December 2024, the country announced a total export ban to the United States on strategic critical minerals including gallium, germanium, and antimony, as well as industrial-strength diamonds and dense synthetic materials, citing national security concerns and responding directly to new US restrictions on semiconductor exports to China.[26] Such efforts are likely to continue into the future.

Geopolitical Risks and Supply Chain Resilience

The extreme concentration of semiconductor manufacturing in Taiwan poses a significant risk to the global supply chain. Political tensions, trade disputes, and potential conflict in the region can disrupt the supply of critical components, impacting industries and economies worldwide. Diversifying supply chains and investing in domestic manufacturing capabilities are essential strategies for building resilience against geopolitical risks.

Initial steps toward this were taken in the passage of the CHIPS Act of 2022, which earmarked $52.7 billion for semiconductor manufacturing, research, and workforce development, plus significant tax credits for private investment in the field. Full implementation of the act has not yet occurred, partly because not enough time has elapsed and partly because the appropriations it called for have not been fully funded.

Investing in domestic manufacturing capabilities and promoting regional cooperation is intended to enhance supply chain security and ensure a steady supply of essential components. For example, TSMC is building facilities in Arizona and Japan, and Samsung is investing in Texas. Japan and India are both investing heavily to modernize and grow their own chipmaking industries, while Southeast Asian countries are focusing on assembly and testing.

Industrial Policy

The US government acquired a 10 percent equity stake in Intel by converting previously awarded but unused government grants proffered under the CHIPS Act into shares.[27] This move aims to support Intel's expansion of domestic chip manufacturing. The government's ownership is passive, with no board seat or governance rights, and includes a warrant for an additional 5 percent if Intel loses majority control of its foundry business.

This deal constitutes an unusual direct investment of the US government in a major private company, possibly signaling a move toward more active government involvement in strategic industries, and a number of analysts have raised concerns about it.[28] They suggest it raises the chance that Intel, or any company in a similar arrangement, would shape its own corporate decision making to align with government or political preferences, compromising its market-driven business priorities. Additionally, it risks distorting competition in markets that would otherwise be free of government stakes, creating potential conflicts between economic efficiency and political objectives. For example, it might lead to an undue bias in favor of meeting national security objectives over maintaining a competitive, innovation-driven product line.

NOTES

1. "Inside an Intel Chip Fab: One of the Cleanest Conference Rooms on Earth," Intel Corporation, March 28, 2018, https://www.intc.com/news-events/press-releases/detail/165/inside-an-intel-chip-fab-one-of-the-cleanest-conference.

2. Anton Shilov, "Former Intel Directors Believe Intel Must Split in Two to Survive," *Tom's Hardware*, October 26, 2024, https://www.tomshardware.com/tech-industry/former-intel-directors-believe-intel-must-split-in-two-to-survive; "Tech Analyst Explains Why Intel Should Be Split Apart and How to Do It," Investing.com, Yahoo Finance, March 5, 2025, https://finance.yahoo.com/news/tech-analyst-explains-why-intel-132943025.html.

3. Jeremy Bowman, "This 1 Number May Ensure TSMC's Market Dominance," *The Motley Fool*, August 17, 2024, https://www.nasdaq.com/articles/1-number-may-ensure-tsmcs-market-dominance.

4. "Global Semiconductor Foundry Revenue Share: Q1 2024," Counterpoint, June 12, 2024, https://www.counterpointresearch.com/insights/global-semiconductor-foundry-market-share/.

5. "Global Semiconductor Foundry Revenue Share," Counterpoint.

6. *2021 State of the U.S. Semiconductor Industry*, Semiconductor Industry Association, 2021, https://www.semiconductors.org/wp-content/uploads/2021/09/2021-SIA-State-of-the-Industry-Report.pdf.

7. "2019 Summit Agenda, Videos, and Slides," Electronics Resurgence Initiative 2.0, Defense Advanced Research Projects Agency, accessed August 30, 2023, https://eri-summit.darpa.mil/2019-archive-keynote-slides.

8. In fact, transistor costs are usually plotted in a semi-log graph, where the log of the cost is plotted against time. In these plots the exponential decline in transistor costs becomes a straight line. The fact that the scale of the y-axis is linear is a clear indication that Moore's law is over.

9. Dion Harris, "Sustainable Strides: How AI and Accelerated Computing Are Driving Energy Efficiency," *NVIDIA Blog*, July 22, 2024, https://blogs.nvidia.com/blog/accelerated-ai-energy-efficiency/.

10. See, for example, Amir Gholami, Zhewei Yao, Sehoon Kim, Coleman Hooper, Michael W. Mahoney, and Kurt Keutzer, "AI and Memory Wall," *IEEE Micro* 44, no. 3 (2024): 33–39, https://doi.org/10.1109/MM.2024.3373763.

11. Mark LaPedus, "SK Hynix Surpasses Samsung in DRAM Share," *Semiecosystem*, Substack, April 11, 2025, https://marklapedus.substack.com/p/sk-hynix-surpasses-samsung-in-dram.

12. Nikkei Asia, "China Makes Inroads in DRAM Chips in Challenge to Samsung and Micron," KrASIA, February 1, 2025, https://kr-asia.com/china-makes-inroads-in-dram-chips-in-challenge-to-samsung-and-micron.

13. Ray Wang, "Mapping China's HBM Advances," November 27, 2024, https://www.chinatalk.media/p/mapping-chinas-hbm-advancement.

14. Paolo Cappaletti and Jon Slaughter, "Embedded Memory Solutions: Charge Storage Based, Resistive and Magnetic," in Andrea Redaelli and Fabio Pellizzer, eds., *Semiconductor Memories and Systems* (Woodhead Publishing, 2022), 159–215, https://doi.org/10.1016/B978-0-12-820758-1.00007-8.

15. Cappaletti and Slaughter, "Embedded Memory Solutions."

16. David A. B. Miller, "Attojoule Optoelectronics for Low-Energy Information Processing and Communications," *Journal of Lightwave Technology* 35, no. 4 (February 2017): 34696, https://doi.org/10.1109/JLT.2017.2647779.

17. Dan Martin and Dan Rosso, "Chipping Away: Assessing and Addressing the Labor Market Gap Facing the U.S. Semiconductor Industry," Semiconductor Industry Association and Oxford Economics, July 25, 2023, https://www.semiconductors.org/chipping-away-assessing-and-addressing-the-labor-market-gap-facing-the-u-s-semiconductor-industry/.

18. Martin and Rosso, "Chipping Away."

19. David Shepardson, "US Commerce Voids Biden's $7.4 Billion Semiconductor Research Grant Deal," Reuters, August 25, 2025, https://www.reuters.com/legal/government/us-commerce-voids-bidens-74-billion-semiconductor-research-grant-deal-2025-08-25/.

20. "Commerce Strengthens Export Controls to Restrict China's Capability to Produce Advanced Semiconductors for Military Applications," news release, Bureau of Industry and Security, US Department of Commerce, December 2, 2024, https://www.bis.gov/press-release/commerce-strengthens-export-controls-restrict-chinas-capability-produce-advanced-semiconductors-military.

21. "Commerce Strengthens Export Controls," Bureau of Industry and Security.

22. "ASML Expects Impact of Updated Export Restrictions to Fall Within Outlook for 2025," news release, ASML, December 2, 2024, https://www.asml.com/en/news/press-releases/2024/asml-expects-impact-of-updated-export-restrictions-to-fall-within-outlook-for-2025.

23. Meg Rithmire and Courtney Han, "The Clean Network and the Future of Global Technology Competition," *Harvard Business Review*, April 2021, https://store.hbr.org/product/the-clean-network-and-the-future-of-global-technology-competition/721045.

24. Associated Press, "U.S. Bans the Sale and Import of Some Tech from Chinese Companies Huawei and ZTE," NPR, November 26, 2022, https://www.npr.org/2022/11/26/1139258274/us-ban-tech-china-huawei-zte.

25. Sujai Shivakumar, Charles Wessner, and Thomas Howell, "Export Controls on U.S. Chip Technology to China," Center for Strategic and International Studies, February 2024.

26. Krystal Bermudez, "China Retaliates Against U.S. Semiconductor Restrictions by Banning Critical Mineral Exports," Foundation for the Defense of Democracies, December 4, 2024, https://www.fdd.org/analysis/policy_briefs/2024/12/04/china-retaliates-against-u-s-semiconductor-restrictions-by-banning-critical-mineral-exports/.

27. "Intel and Trump Administration Reach Historic Agreement to Accelerate American Technology and Manufacturing Leadership," news release, Intel Corporation, August 22, 2025, https://www.intc.com/news-events/press-releases/detail/1748/intel-and-trump-administration-reach-historic-agreement-to.

28. See, for example, Ross Kerber, "Investors Worry Trump's Intel Deal Kicks Off Era of US Industrial Policy," Reuters, August 27, 2025, https://www.reuters.com/sustainability/boards-policy-regulation/investors-worry-trumps-intel-deal-kicks-off-era-us-industrial-policy-2025-08-27/; Alexander Bolton, "Donald Trump's Government Buying Stake in Intel Prompts GOP Complaints," *The Hill*, August 26, 2025, https://thehill.com/homenews/senate/5469740-gop-criticizes-trump-intel-deal/; David Shepardson and Arsheeya Bajwa, "Intel Warns US Stake Could Hurt International Sales, Future Grants," China, Reuters, August 25, 2025, https://www.reuters.com/world/china/intel-warns-us-stake-could-hurt-international-sales-future-grants-2025-08-25/.

STANFORD EXPERT CONTRIBUTORS

Dr. Mark A. Horowitz
SETR Faculty Council, Fortinet Founders Chair of the Department of Electrical Engineering, Yahoo! Founders Professor in the School of Engineering, and Professor of Computer Science

Dr. David Miller
W. M. Keck Foundation Professor of Electrical Engineering and Professor, by courtesy, of Applied Physics

Dr. Asir Intisar Khan
SETR Fellow and Visiting Postdoctoral Scholar in Electrical Engineering

10

SPACE

KEY TAKEAWAYS

- A burgeoning "NewSpace" economy driven by private innovation and investment is transforming space launch, in-space logistics, communications, and key space actors in a domain that until now has been dominated by superpower governments.

- Space is a finite planetary resource. Because of dramatic increases in satellites, debris, and geopolitical space competition, new technologies and new international policy frameworks will be needed to manage the traffic of vehicles, prevent international conflict in space, and ensure responsible stewardship of this global commons.

- The Trump administration has shifted priorities heavily toward human exploration of the Moon and Mars. This is at the expense of robotic exploration, space science, and aeronautics missions, leading to significant planned budget and personnel cuts to NASA. This trend may risk the long-term superiority of the United States in the global race for talent and technology.

Overview

Sputnik 1 was the world's first artificial satellite, a technology demonstration placed into orbit by the Soviet Union in 1957. Sixty-nine years later, humankind operates many thousands of satellites to provide communications, navigation, and Earth observation imagery relied upon in many walks of life. A substantial amount of scientific discovery is also made possible with space-borne instrumentation. Additionally, space operations support military forces on Earth, and thus space itself is a domain in which international conflict and competition play out.

Today, the global space economy is growing at about 7 percent per annum.[1] Valued at $600 billion in 2024, it is forecast to potentially reach $1.8 trillion by 2035.[2] This growth is driven by space-based technologies and their impacts on various industries, including defense, transportation, and consumer goods. One distinctive feature of the space industry is the predominance of government investment, which totaled $135 billion globally in 2024, more than $75 billion of which came from the United States.[3] While ownership of space assets is gradually shifting from governments to private providers, private-sector investment has declined for three consecutive years—from its all-time high of $18 billion in 2021 to $5.9 billion in 2024—reflecting ongoing challenges in commercialization and intensifying global competition.[4] The number of satellites launched per year has grown at a cumulative annual rate of over 50 percent from 2019 to 2024, supported by an increase in global rocket launches (271 in 2024[5]).

At its core, a space mission includes four components:

- The mission objectives, which can be scientific, commercial, military, or a combination thereof
- A space segment, which includes the spacecraft and the orbits that have been selected to accomplish the mission objectives
- A ground segment, which includes the rocket launcher, ground stations, and mission control centers
- A user segment, which includes all the users and stakeholders of the space mission

Space systems can be categorized in various ways. One such sorting factor is whether they are crewed or uncrewed. The International Space Station (ISS) is currently the nexus for spaceflight; since 2011, US-crewed access to the ISS has been via rockets operated by Russia—and more recently through vehicles provided by SpaceX and Boeing. In the future, the NASA-operated Artemis program plans to launch its first crewed mission, a Moon flyby, in early 2026, followed by a Moon landing in 2027 or 2028. Uncrewed systems include those for Earth and planetary remote sensing (such as Planet Labs' Dove satellites); communication and navigation (such as the United States' Global Positioning System, or GPS, satellites); astronomy and astrophysics (such as the James Webb Space

FIGURE 10.1 Fires and damage at Antonov Airport in Ukraine, as seen from a commercial satellite constellation

Source: © 2022 Maxar Technologies

Telescope); space logistics and in-space assembly and manufacturing (such as Northrop Grumman's Mission Extension Vehicles, or MEVs); and planetary exploration (such as the Mars Perseverance rover).

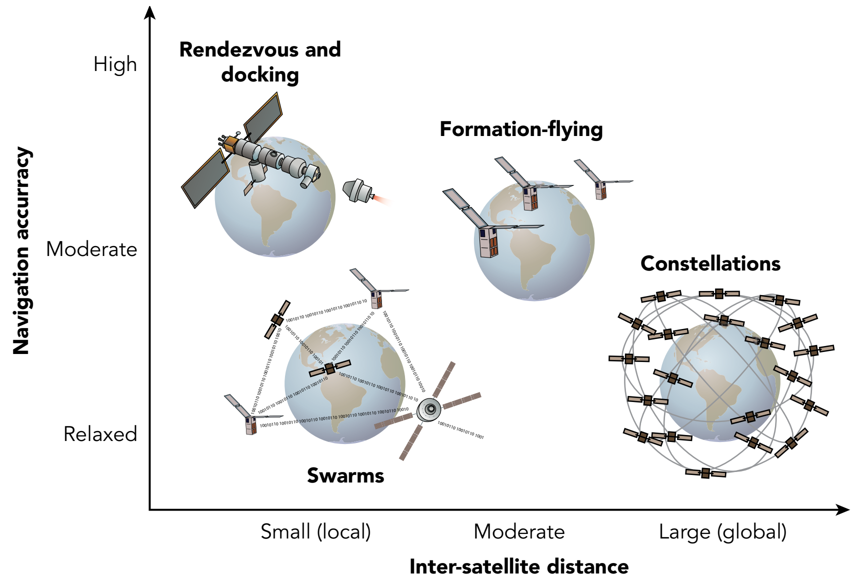
Space systems can also be characterized by size. Very large structures include the ISS, whose mass is 420 tons, and proposed future space stations that are part of the NASA-funded Commercial LEO Destinations (CLD) program, such as Orbital Reef (Blue Origin), and Starlab (Starlab Space, a joint venture of Voyager Space and Airbus). Vastly smaller satellites, called smallsats, weigh under 500 kilograms (kg).[6] CubeSats are the most popular smallsat format, with each CubeSat unit measuring 10 by 10 by 10 centimeters (cm) and with a mass of 1.33 kg (a couple of pounds). They can also be combined to build larger satellites. CubeSats support a growing commercial market, providing communications, Earth imagery, and other capabilities. Today, a large majority of functional satellites weigh between 100 and 1,000 kg.

Another classification of space systems reflects their trajectories. For example, objects in orbit around Earth can be in low Earth orbit (LEO), which is less than 1,000 kilometers (km) in altitude; medium Earth orbit (MEO), which is between 1,000 and 35,000 km in altitude; high elliptical orbit; and geosynchronous orbit (GEO), with an orbit period equivalent to one Earth day. The image in figure 10.1 was obtained by a Maxar Technologies commercial satellite in LEO.

A further categorization of space systems focuses on their composition. Distributed space systems, comprising multiple interacting spacecraft, can achieve objectives that are difficult or impossible for a single spacecraft. These systems take various architectural forms (see figure 10.2), defined by parameters like inter-spacecraft distances, required navigational accuracy, and number of satellites. They contrast with traditional single-spacecraft systems and offer expanded capabilities in space operations. Different compositions of space systems include the following:

- Constellations separated by tens of thousands of kilometers so that they may provide global

FIGURE 10.2 Characterizing distributed space systems



Source: Adapted from a diagram by Simone D'Amico

FIGURE 10.3 An artist's conception of the NASA Starling satellite swarm in space

Source: NASA / Blue Canyon Technologies

coverage for navigation, communications, and remote sensing services.

- Rendezvous and docking to support crew transportation, removal of space debris from orbit, in-orbit servicing of satellites, and assembly of larger structures in space. This involves small separations and high positional accuracy.
- Formation-flying architectures for observational missions that call for large effective apertures, such as space-based telescopes whose optical components are controlled very precisely at separations of tens to hundreds of meters.
- Swarms that cooperatively sense the environment or share resources such as power, computation, and communications but whose components do not necessarily need to be at fixed distances from one another (see figure 10.3).

Key Developments

Impacts of Space Technologies

Space technologies have proven their value to the national interest. Some of the most important applications today include the following:

Navigation This includes positioning, navigation, and timing (PNT) services around the world and in space. GPS and similar services operated by other nations help people know their position and velocity, whether on land, on the ocean surface, in the air, or in space. Less well known is the timing information that GPS provides—timing that is accurate to the nanosecond available anywhere in the world. This is a key tool for the financial sector, electric power grid, and transportation. Companies such as Xona Space

Systems, a start-up founded by Stanford alumni, have begun developing GPS alternatives that aim to deliver even greater precision and robustness. More recently, driven by the explosion in the number of satellites and other spacecraft in orbit, interest has been growing quickly in characterizing and managing in-space objects.

Communications Satellites provide vital communications in remote areas and for mobile users, complementing the terrestrial networks that carry most long-haul communications. Companies like SpaceX's Starlink, Amazon's Project Kuiper, Eutelsat OneWeb, and Astranis aim to offer low-latency, wide-coverage satellite internet. Recent innovations include optical communication systems, which use light for higher bandwidth and security. Vertical integration across orbital, aerial, marine, and ground segments will be a key asset for future communication infrastructure. Several start-ups, such as Aalyria and SpiderOak, are also developing technologies for the orchestration of different networks to ensure robust and secure communications.

Remote sensing Remote sensing satellites, with their unique vantage point and sophisticated sensors, can rapidly gather extensive data about areas and objects of interest. These data are then integrated and used to train artificial intelligence (AI) models to create a "digital twin" of Earth, enhancing prediction and simulation of terrestrial phenomena and responses to them. Applications include disaster response, environmental monitoring, topographical mapping, and geospatial intelligence tracking human, animal, and marine activity. Governments are expanding remote sensing programs, complemented by commercial companies like BlackSky, Maxar, Planet Labs, Spire Global, and ICEYE. Recent efforts have focused on increasing data resolution, reducing response times, and exploring other valuable information modes such as hyperspectral imaging,[7] synthetic aperture radar,[8] and radio-frequency sounding (exploration of the environment through the use and exploitation of radio waves).[9]

Scientific research Space-based astronomy and exploration provide in-depth insights into the origins of planets, stars, galaxies, and life on Earth. The past few years have seen significant strides in solar system exploration, particularly involving asteroids. NASA's OSIRIS-REx mission successfully returned asteroid samples to Earth in September 2023, while the NASA Psyche mission was launched in October 2023 and is en route to examine at close range a metal-rich asteroid worth potentially quadrillions of dollars.

Space transportation The space transportation industry has seen launch costs drop by more than an order of magnitude over a couple of decades to $1,500 per kilogram in 2021.[10] Companies like SpaceX, Rocket Lab, Blue Origin, Stoke Space, and Virgin Galactic have made progress in providing reliable launches and developing new vehicles. SpaceX's Starship—the most powerful rocket ever built (see figure 10.4)—could dramatically reduce the cost of achieving LEO, aspirationally making this between 10 and 100 times cheaper than today.[11] Also, in January 2025, Blue Origin launched its reusable high-volume, heavy-lift New Glenn rocket, which successfully deployed a prototype of its Blue Ring platform in MEO.[12] (Blue Ring can serve as a satellite support platform and a space tug that transfers cargo between different orbits.) While SpaceX currently dominates the space launch sector, overreliance on a single company could prove risky for the US government, especially given the multiple failed test launches of Starship in 2025.

Meanwhile, Blue Origin, Voyager Space, Axiom Space, and Vast are developing commercial space stations to replace the ISS, which NASA plans to decommission in 2030. These new stations aim to ensure continued orbital research and expand human presence in space.

National security Spacecraft constantly scan Earth for launches of ballistic and hypersonic missiles aimed at the United States or its allies, nuclear weapons tests anywhere in the world, radio traffic and radar

FIGURE 10.4 SpaceX's Starship could dramatically reduce the cost of achieving LEO

Source: SpaceX, CC BY-NC 2.0

signals from other countries, and the movements of allies and enemies in military contexts. A major focus of the Trump administration is the "Golden Dome," a proposed multilayer system partly based in space used for threat monitoring and missile defense.[13] US government investment in space for national security purposes continues to grow, including new commercial partnerships focused on data sharing for tracking objects in space and on Earth, satellite internet for battlefield communications, and in-space logistics (inspection and servicing) to maintain space superiority and safety.

Trends in Space Technology

Privatization, miniaturization, and reusability The space sector is shifting from government-owned legacy systems and their long development timelines and mission lifetimes to a "NewSpace" economy driven by private companies. This privatization makes space technologies more accessible and less expensive. CubeSats and reusable rockets like SpaceX's Falcon 9 exemplify private-sector innovations enabling new opportunities. Governments are also embracing small spacecraft and on-demand launches to expand space capabilities cost-effectively. The combination of smallsats and distributed architecture (e.g., constellations) offers advantages in reduced costs, faster development timelines, frequent technology updates, and improved resilience, flexibility, and performance.

However, the private sector's rapidly increasing role in space also presents new challenges. These include dealing with risks inherent in dual-use space technologies (for example, adversaries could use technology that was designed for removing space debris to attack other satellites); managing crises in a realm where lines separating individual private actors, the space sector as a whole, and government actors are increasingly blurred; differentiating between accidents and malevolent actions; and relying on companies whose interests may not be fully aligned with those of the US government.

The Moon rush Recent years have seen a renewed desire to maintain a permanent human presence in lunar orbit and on the lunar surface. The abundance of certain materials on the Moon provides opportunities for mining and manufacturing, known as in-situ resource utilization. Such activities would reduce the amounts of material that would otherwise have to be transported from Earth. Combined with the significantly lower amount of fuel needed to launch from the Moon rather than from Earth, moon mining and manufacturing facilitates the construction of moon bases, the conduct of space exploration missions, and even launches into LEO that could be undertaken with hardware manufactured with materials from the Moon.

There have been a number of successful lunar landings recently from both the commercial and the civil sectors, including the United States, China, India, and Japan.[14] The NASA-led Artemis program is developing a new launch system, lunar-orbiting space station, lunar base camp, and lunar terrain vehicles, among other things, as steps needed for establishing a permanent human presence on the Moon.[15]

Over the Horizon

Advances in Small Satellite Technology

NASA has identified a list of issues that are restraining growth in usage of small spacecraft.[16] They include limitations in launch capacity, autonomous capabilities, PNT capabilities, and propulsion systems. The past year has seen several mission failures due to these and other shortfalls.

NASA is responding to these challenges via technology demonstration missions such as Starling, which, in 2024, became the first successful in-orbit example of several critical autonomous swarming technologies.[17] Among Starling's payloads is an optical PNT system newly developed by Stanford's Space Rendezvous Laboratory. This applies onboard cameras and advanced algorithms to not only navigate multiple satellites cooperatively, but also to characterize resident space objects (RSO)—any human-made or natural object orbiting Earth—using only visual data. In doing so, it addresses a key technological gap for small spacecraft, which must typically rely on jammable GPS or expensive ground-based resources for navigation. However, more work is needed to take full advantage of smallsat architectures—and, by extension, distributed architectures featuring many smallsats working together.

New Applications of Space Technologies

Manufacturing For certain types of manufacturing, such as specialized pharmaceuticals, optics, and semiconductors, space offers two major advantages over terrestrial manufacturing. Because the vacuum of space is very clean, minimizing contamination is much easier. Further, space's microgravity environment means that phenomena resulting from the effects of gravity—such as sedimentation, buoyancy, thermal convection, and hydrostatic pressure—can be minimized. This enables, for example, the fabrication of more perfect crystals and more perfect shapes. Production processes for biological materials, medicines, metallizations, polymers, semiconductors, and electronics may benefit.

Mining The Moon and asteroids may well have vast storehouses of useful minerals that are hard to find or extract on Earth, such as rare-earth elements used in batteries and catalytic converters as well as in guidance systems and other defense applications. Helium-3 found on the Moon may be an important source of fuel for nuclear fusion reactors. Future space-mining operations could bring these resources back to Earth to meet growing demand in a sustainable way. Mining of regolith (loose rock that sits atop bedrock) and ice on the Moon is also critical for enabling a permanent human presence there and supporting subsequent expansion into the solar system.

Power generation Most orbits are exposed to a constant and intense sunlight, which can be a potential source of clean energy generation. Initial technology demonstrations have been performed in the past two years,[18] and several companies are trying to unlock the potential in this area through two main approaches:

- Solar panels in space that can capture solar energy, convert it to microwaves or laser light, and beam it down to Earth, where dedicated ground stations receive the transmitted energy
- Large mirrors in space that can directly reflect sunlight to existing solar farms on the ground

Both approaches can be advantageous for areas on Earth that cannot easily receive power or fuel supplies or access these around the clock. While this technology is still at a very early stage, interest in it has been growing steadily.

Space situational awareness (SSA) The number of active satellites has increased from roughly 1,000 in 2014 to about 11,000 in 2025—a figure that will likely rise to several tens of thousands in the next decade. In addition, the European Space Agency estimates that about 1.2 million pieces of debris larger than 1 cm in size are in orbit, many of which are dangerous to satellites and space stations.[19] Some potential methods of tracking this ultrasmall debris by leveraging its electrical charge and the plasma environment of space have been proposed.[20]

Traditionally, military organizations such as the North American Aerospace Defense Command and the US Space Force (USSF) were responsible for tracking space objects. However, beginning in fall 2024,[21] the US Department of Commerce's Traffic Coordination System for Space (TraCSS) program began taking over civil and commercial SSA responsibilities from the Department of Defense. TraCSS provides basic SSA data and services to support the safety of civil and private spaceflight operations.[22]

This transition was motivated by the fact that SSA is important to both the civil and commercial sectors and to both national and international actors, not only to the US defense sector. Furthermore, this transition will alleviate the burden on the USSF, allowing it to focus on national security priorities in space. Commercial players will be heavily involved in providing a more accurate and responsive space-traffic management system by improving ground tracking capabilities using radars and telescopes, and deploying satellites and payloads for more timely and responsive on-orbit surveillance.

In-space servicing, assembly, and manufacturing (ISAM) Leadership, security, and sustainability in space require ISAM capabilities to approach, inspect, repair, refuel, or remove space assets without jeopardizing the space environment.[23] Spacecraft autonomy, in combination with rendezvous, proximity operations, and docking (RPOD), is a critical technology for ISAM. RPOD refers to the ability of spacecraft to operate autonomously in combination with the ability to approach one another precisely and conduct close-up operations. Despite significant interest, only a handful of missions have demonstrated early RPOD capabilities in orbit; these include recent successes achieved by Astroscale (using its ADRAS-J smallsat in LEO) and Northrop Grumman (via its larger MEV satellites in GEO).

Exploration A critical limitation for space exploration missions is travel time: Getting to the outer solar system can take ten years or more. As spacecraft fly ever farther from the Sun, they will need novel forms of power, such as sources driven by nuclear reactions, for the propulsive energy needed to make their missions possible.[24] Better propulsion systems that can be quickly deployed will also be needed to intercept interstellar objects so that samples can be collected from them.

On-demand space exploration missions Today, it takes a very long time to prepare for exploration missions, which means that targets of opportunity that suddenly appear cannot be visited by such missions. An on-demand capability would enable the close-up investigation of suddenly appearing targets such as the Oumuamua interstellar object, which passed through the solar system in 2017. Undertaking such a mission requires that a spacecraft can be made ready for launch shortly after the target is identified. Because such targets are likely to originate outside our solar system, the scientific return from bringing back a sample would be enormous.

Policy Issues

Shift in US Executive Branch Priorities

The Trump administration has been reformulating national space priorities. While the current and proposed budgets for the domain aim to bolster human exploration to the Moon and Mars, they have also imposed significant cuts to numerous ongoing and planned scientific missions related to interplanetary exploration, aviation, and space science. These cuts have led to an outflow of talent from NASA, with the loss of four thousand employees (a fifth of the workforce) already.[25] Similar budget cuts have been made to the National Science Foundation, Department of Energy, and other space-adjacent federal agencies. These cuts have a large impact not only on academic research and development, but also on private companies, which receive a large amount of federal funding to serve as contractors for government space programs. These policies may put the United States at a competitive disadvantage relative to China and other global powers.

The Grand Challenge of Sustainability

Sustainability encompasses both terrestrial sustainability enabled by space and the sustainability of humankind's use of space.

- **Sustainability enabled by space** incorporates several of the technologies described above: for example, creating Earth's digital twin for disaster prevention and management, which requires integrating data from industry, government, and academia with advanced machine learning techniques. Space-based solar power and resource extraction from the Moon and other celestial bodies, among other facets, illustrate the potential here.

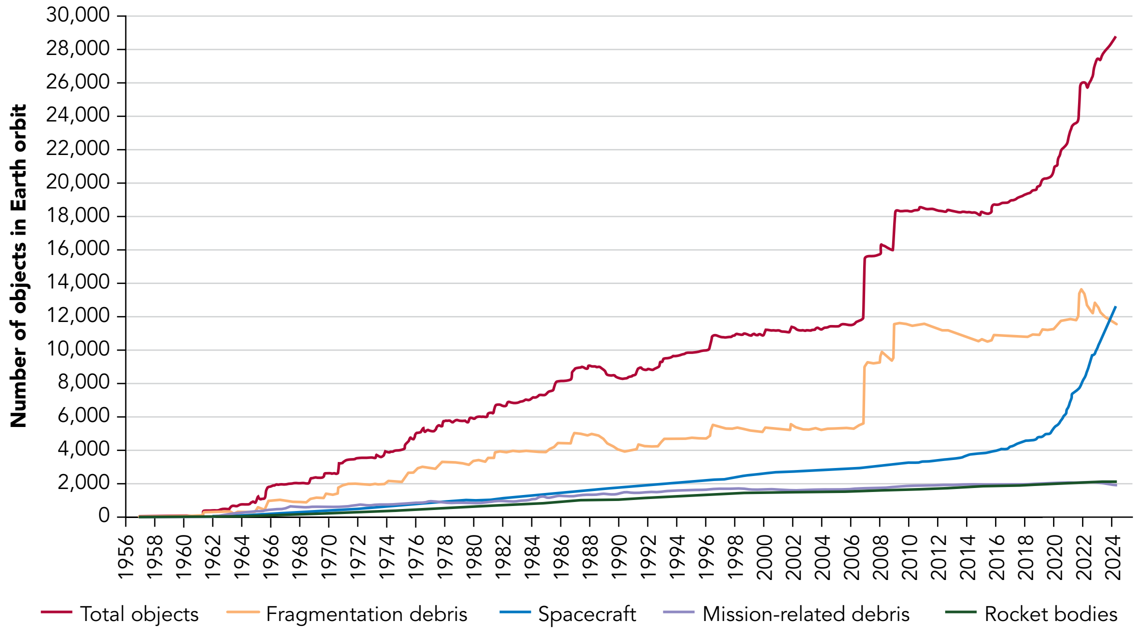
- **Sustainability of space** aims to create a circular, equitable space economy. Unlike Earth's organized transportation systems (which include traffic laws and gas stations), space lacks similar infrastructure. Addressing this requires making space assets reusable, establishing orbital services, managing debris, and quantifying orbital capacity. Space traffic management is essential to handle the increasing number of assets around the Earth and Moon. Developing guidelines for fair and safe orbital behavior—which don't currently exist—is essential.

The world ultimately faces a spaceflight sustainability paradox: The growing use of space to support sustainability and security on Earth will lead to more adverse impacts on the space environment itself. For example, multiple constellations of remote sensing satellites will contribute to greater space traffic challenges. Managing this complex issue will require advances in both policy and technology.

Triple-Helix Innovation

Collaborative efforts between academia and industry often focus on technology commercialization and real-world demonstrations, frequently supported by governments. This cooperative model, known as triple-helix innovation, combines academia, industry, and government. Notable examples include the proposed $2 billion Berkeley Space Center collaboration with NASA's Ames Research Center and Stanford University's Center for AEroSpace

FIGURE 10.5 The number of tracked space objects larger than 10 centimeters has grown rapidly



Source: Adapted from *Orbital Debris Quarterly News* 28, no. 3 (July 2024): 10

Autonomy Research (CAESAR), which focuses on AI-driven autonomy with Blue Origin, Redwire, and government agencies.

Space Governance

International and national space governance has not developed at the same rapid pace as space technology. Existing legal frameworks—many of which are products of the Cold War—do not address wide swaths of current activities and are often contested in scope and interpretation.[26] Attempts at improvement have often stagnated due to differing geopolitical aims. Even within the United States, space assets are not designated as critical infrastructure by the government despite their importance, and growth in space activity far outpaces the capabilities of current licensing processes run by the Federal Aviation Administration and the Federal Communications Commission (FCC).

Nonetheless, a number of developments in the past couple of years are notable. NASA released its strategy, including actionable objectives, for sustainability in space activities in Earth orbit.[27] It also promised to release similar strategies in the future for activities on Earth; for the orbital area near and around the Moon known as cislunar space; and for deep space, including other celestial bodies. In addition, the FCC issued its first-ever fine for a satellite not properly disposed of from geostationary orbit.[28] These short-term policy advances must be unified with a longer-term vision encompassing the next fifty to one hundred years to effectively address national security needs, support the space industry's continued development, and realize the responsible use of space as a global commons.

Maintaining Space Access

The number of objects in space has grown rapidly. Figure 10.5 shows the total number of tracked space

While prestige remains a factor, the current [Moon] race focuses on establishing a lunar presence for strategic and economic advantages.

objects larger than 10 cm since 1959. Today, there are nearly 30,000 such objects, about 10,000 of which are working satellites. There are also an estimated 1.1 million fragments between 1 and 10 cm in size.[29] With so many objects in space, the risks of collision between them are growing. Each collision has the capacity to create even more debris, leading to a catastrophic chain reaction known as the Kessler syndrome, which would effectively block access to space. In addition, increasing volumes of space traffic (future mega-constellations will consist of tens of thousands of satellites) may lead to communications interference, and coordination of space activities such as orbit planning will be increasingly difficult to manage.

To tackle this issue, new domestic safety legislation and international cooperation will be needed for accurate tracking of space objects, facilitating the use of automated collision-avoidance systems, and removing debris from orbit. Similarly, more consistent guidelines will be needed to govern behavior in space, how space operations are conducted, and the sharing of data for situational awareness. Transparency and coordination among all players will be key, and the United States is in a good position to take a leading role among like-minded nations in advocating for these kinds of changes in space access.

Geopolitics, Security, and Conflict in Space

Many issues arise with respect to space and geopolitics. A key example is the Outer Space Treaty (OST), which entered into force in 1967; today, 117 countries are parties to the treaty, and 22 more have signed but not ratified it. Among other things, the treaty prohibits the placement of nuclear weapons or other weapons of mass destruction in space.

Recent evidence suggests the OST's norms are eroding. In 2024 Russia vetoed a United Nations resolution prohibiting the deployment of nuclear weapons in space, despite being a party to the OST; senior US officials revealed that they believe Russia is developing a satellite to carry nuclear weapons into LEO, where a detonation could destroy all satellite activity there for up to a year.[30] In addition, there is no treaty, OST or otherwise, that limits other military uses of space, including the placement of conventional weapons in orbit.

A second issue relates to nonnuclear anti-satellite weaponry and capabilities. To date, four nations—China, Russia, India, and the United States—have successfully tested kinetic anti-satellite weapons capable of physically destroying satellites in space. (Every such test has produced a significant amount of space debris.) More broadly, countries are developing a range of capabilities, from the ground and in space, to degrade, deny, and even destroy satellites of other nations. Cyberattacks are an important element of the non-kinetic threat spectrum against space missions, which can lead to data corruption, jamming, and hijacking of space intelligence providers and customers.[31]

A third issue involves various national efforts to reach the Moon. To facilitate an orderly and peaceful exploitation of the Moon's resources that is consistent with the OST, NASA and its partners have also proposed the Artemis Accords, which define "principles for cooperation in the civil exploration and use of the Moon, Mars, comets and asteroids for peaceful purposes."[32] So far, forty-three nations have signed the accords—but notably not Russia or China, which are among the parties seeking to establish a permanent Moon presence.

Nations today are engaged in a new "race to the Moon," though with different motivations than in the 1960s. While prestige remains a factor, the current race focuses on establishing a lunar presence for strategic and economic advantages. The first nation to establish a lunar presence successfully may well gain a first-mover advantage that enables it to be in a stronger position to set the terms for others to come. Although the OST prohibits claiming lunar sovereignty, there are concerns that nations might disregard this for national interests.[33] The possibility of a nation taking military action to prevent others from establishing their own lunar presence highlights the potential for conflict in this new space race.

Finally, in the past couple of years, the rise of the private sector's importance in providing capabilities for rapid space launch and space-based communications has been dominated by SpaceX and Starlink, which are owned by the same person. In 2022, the CEO of Starlink, which Ukrainian military forces relied on for communications, denied its use to conduct military operations around Sevastopol, in Crimea—thus directly interfering with the execution of Ukrainian battle plans.[34] In September 2024, NASA turned to SpaceX to return to Earth two US astronauts left on the ISS when their Boeing-built Starliner spacecraft experienced operational failures and was brought to Earth without them.

Such incidents demonstrate the extreme dependence of the US government on capabilities provided by a very limited number of companies and raise important policy questions of how to ensure that US space efforts can continue in accordance with US national interests.

NOTES

1. "Next in Space 2025," PwC, April 3, 2025, https://www.pwc.com/us/en/industries/industrial-products/library/space-industry-trends.html.

2. *Space Economy Report*, 11th ed., NovaSpace, 2025, https://nova.space/hub/product/space-economy-report/.

3. *Government Space Programs*, 24th ed., NovaSpace, 2025, https://nova.space/hub/product/government-space-programs/.

4. *Highlights of the 2024 Space Economy*, NovaSpace, 2025, https://nova.space/in-the-loop/highlights-of-the-2024-space-economy/.

5. "2024 Rocket Launch Recap," RocketLaunch.org, accessed July 5, 2025, https://rocketlaunch.org/rocket-launch-recap/2024.

6. Stevan M. Spremo, Alan R. Crocker, and Tina L. Panontin, *Small Spacecraft Overview*, NASA Technical Reports Server, May 15, 2017, https://ntrs.nasa.gov/api/citations/20190031730.

7. Anuja Bhargava, Ashish Sachdeva, Kulbushan Sharma, Mohammed H. Alsharif, Peerapong Uthansakul, and Monthippa Uthansakul, "Hyperspectral Imaging and Its Applications: A Review," *Heliyon* 10, no. 12 (2024): e33208, https://doi.org/10.1016/j.heliyon.2024.e33208.

8. Kelsey Herndon, Franz Meyer, Africa Flores, Emil Cherrington, and Leah Kucera, "What Is Synthetic Aperture Radar?," NASA Earthdata, accessed September 13, 2024, https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar.

9. J. Wickert, T. Schmidt, M. Schmidt, L. Sanchez, and M. Sehnal, "GNSS Radio Occultation," Global Geodetic Observing System, accessed September 13, 2024, https://ggos.org/item/gnss-radio-occultation/.

10. Ryan Brukardt, "How Will the Space Economy Change the World?," McKinsey & Company, November 28, 2022, https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/how-will-the-space-economy-change-the-world.

11. Citi GPS, "Space: The Dawn of a New Age," Citi, May 9, 2022, https://www.citigroup.com/global/insights/space_20220509.

12. "Blue Origin's New Glenn Reaches Orbit," Blue Origin, January 16, 2025, https://www.blueorigin.com/news/new-glenn-ng-1-mission.

13. Geoff Brumfiel, "Trump Wants a Golden Dome over America. Here's What It Would Take," NPR, April 22, 2025, https://www.npr.org/2025/04/22/g-s1-61658/trump-golden-dome-america-iron-military-defense.

14. The landing was a mixed success. "Intuitive Machines Calls Moon Mission a Success Despite Soft Crash Landing," *Space & Defense*, February 26, 2024, https://spaceanddefense.io/intuitive-machines-calls-moon-mission-a-success-despite-soft-crash-landing/.

15. "Artemis Plan: NASA's Lunar Exploration Program Overview," NASA, September 2020, https://www.nasa.gov/wp-content/upaloads/2020/12/artemis_plan-20200921.pdf?emrc=f43185.

16. Ames Research Center, "State-of-the-Art Small Spacecraft Technology," National Aeronautics and Space Administration, February 2024, https://www.nasa.gov/smallsat-institute/sst-soa/.

17. Justin Kruger, Simone D'Amico, and Soon S. Hwang, "Starling Formation-Flying Optical Experiment: Initial Operations and Flight Results," paper presented at the 38th Small Satellite Conference, Logan, UT, August 8, 2024, https://digitalcommons.usu.edu/smallsat/2024/all2024/130/.

18. “In a First, Caltech’s Space Solar Power Demonstrator Wirelessly Transmits Power in Space,” California Institute of Technology, June 1, 2023, https://www.caltech.edu/about/news/in-a-first-caltechs-space-solar-power-demonstrator-wirelessly-transmits-power-in-space.

19. “ESA Space Environment Report 2025,” European Space Agency, January 4, 2025, https://www.esa.int/Space_Safety/Space_Debris/ESA_Space_Environment_Report_2025.

20. Abhijit Sen, Rupak Mukherjee, Sharad K. Yadav, Chris Crabtree, and Gurudas Ganguli, “Electromagnetic Pinned Solutions for Space Debris Detection,” *Physics of Plasmas* 30, no. 1 (2023): 012301, https://doi.org/10.1063/5.0099201.

21. “Space Commerce Highlights,” Office of Space Commerce, US Department of Commerce, September 2024, https://space.commerce.gov/wp-content/uploads/Space-Commerce-Highlights-OSC-Newsletter-September-2024.pdf.

22. Traffic Coordination System for Space (TraCSS), Office of Space Commerce, Department of Commerce, September 2024, https://space.commerce.gov/traffic-coordination-system-for-space-tracss/.

23. “White House Office of Science and Technology Policy Unveils National In-Space Servicing, Assembly, and Manufacturing (ISAM) Implementation Plan,” news release, The White House, December 16, 2022, https://www.whitehouse.gov/ostp/news-updates/2022/12/16/white-house-office-of-science-and-technology-policy-unveils-national-in-space-servicing-assembly-and-manufacturing-isam-implementation-plan.

24. Beth Ridgeway, ed., “Space Nuclear Propulsion,” NASA, last modified August 19, 2024, https://www.nasa.gov/tdm/space-nuclear-propulsion.

25. Mike Wall, “NASA Losing Nearly 4,000 Employees to Trump Administration’s ‘Deferred Resignation’ Program,” Space.com, July 25, 2025, https://www.space.com/space-exploration/nasa-losing-nearly-4-000-employees-to-trump-administrations-deferred-resignation-program.

26. Sophie Goguichvili, Alan Linenberger, Amber Gillette, and Alexandra Novak, “The Global Legal Landscape of Space: Who Writes the Rules on the Final Frontier?,” Wilson Center, October 1, 2021, https://www.wilsoncenter.org/article/global-legal-landscape-space-who-writes-rules-final-frontier.

27. “NASA’s Space Sustainability Strategy: Volume 1: Earth Orbit,” NASA, April 9, 2024, https://www.nasa.gov/wp-content/uploads/2024/04/nasa-space-sustainability-strategy-march-20-2024-tagged3.pdf.

28. “FCC Takes First Space Debris Enforcement Action,” news release, Space Bureau, Enforcement Bureau, and Media Relations Bureau, Federal Communications Commission, October 2, 2023, https://www.fcc.gov/document/fcc-takes-first-space-debris-enforcement-action.

29. “Space Debris by the Numbers,” Space Debris Office, European Space Agency, last modified August 15, 2024, https://www.esa.int/Space_Safety/Space_Debris/Space_debris_by_the_numbers.

30. Audrey Decker, “Russian Space Nuke Could Render Low-Earth Orbit Unusable for a Year, US Official Says,” *Defense One*, May 1, 2024, https://www.defenseone.com/threats/2024/05/russian-space-nuke-could-render-low-earth-orbit-unusable-year-us-official-says/396245/; Robert Wood, “Remarks at a UN General Assembly Meeting on Russia’s Veto of the U.S. and Japan-drafted UNSC Resolution on Preventing Nuclear Weapons in Outer Space,” US Mission to the United Nations, May 6, 2024, https://usun.usmission.gov/remarks-at-a-un-general-assembly-meeting-on-russias-veto-of-the-u-s-and-japan-drafted-unsc-resolution-on-preventing-nuclear-weapons-in-outer-s/.

31. *Global Space Economy*, 3rd ed., Northern Sky Research, January 2023, https://www.nsr.com/?research=global-space-economy-3rd-edition.

32. “The Artemis Accords,” NASA, accessed September 13, 2024, https://www.nasa.gov/artemis-accords.

33. China has chosen to disregard the findings of a tribunal convened pursuant to the United Nations Convention on the Law of the Sea (UNCLOS) that ruled against Chinese claims regarding its land reclamation activities in the South China Sea, despite the fact that China is a formal signatory to UNCLOS. See In the Matter of the South China Sea Arbitration (Phil. v. China), Case No. 2013-19, Award of July 12, 2016, ITLOS Rep., https://pcacases.com/web/sendAttach/2086.

34. Wes J. Bryant, “When a CEO Plays President: Musk, Starlink, and the War in Ukraine,” *Irregular Warfare Initiative*, October 17, 2023, https://irregularwarfare.org/articles/when-a-ceo-plays-president-musk-starlink-and-the-war-in-ukraine/.

STANFORD EXPERT CONTRIBUTORS

Dr. Simone D’Amico
SETR Faculty Council and Associate Professor of Aeronautics and Astronautics and, by courtesy, of Geophysics

Dr. Sigrid Elschot
Professor of Aeronautics and Astronautics

Dr. Anton Ermakov
Assistant Professor of Aeronautics and Astronautics and, by courtesy, of Geophysics and of Earth and Planetary Sciences

Dr. Debbie Senesky
Associate Professor of Aeronautics and Astronautics and of Electrical Engineering

Walter J. Manuel
SETR Fellow and PhD Candidate in Aeronautics and Astronautics

Yuji Takubo
SETR Fellow and PhD Candidate in Aeronautics and Astronautics

11

CROSSCUTTING THEMES AND COMMONALITIES

None of the individual technology areas covered in chapters 1 through 10 operate in a vacuum. It is crucial that policymakers consider broader, crosscutting themes that influence how technology progresses over time as well as the key common drivers that can accelerate or hinder progress. By devoting an entire chapter to them, we wish to underline the important similarities in how people and institutions make progress and emphasize that, when crafting policy for individual domains, it is essential to take a holistic view of the emerging tech landscape and the factors affecting it.

This chapter organizes crosscutting themes into four categories:

- **Governance and Geopolitics of Emerging Technology** examines how governments and political systems shape global technological progress.
- **Innovation Pathways and Patterns of Progress** explores the diverse ways in which technological progress unfolds.
- **Human Capital and Knowledge Ecosystems** highlights the critical roles of people, universities, and funding structures in driving and sustaining innovation.
- **Infrastructure for Innovation** encompasses vital systems and structures that support innovation on a large scale.

Governance and Geopolitics of Emerging Technology

KEY TAKEAWAYS

- Innovation that emerges too fast threatens the legitimate interests of those who might be negatively affected, while innovation that moves too slowly increases the likelihood that a nation will lose first-mover advantages.
- National monopolies on technology are increasingly difficult to maintain. Even innovations that are solely American born (an increasingly rare occurrence) are unlikely to remain in the exclusive control of American actors for long periods.
- The US government is no longer the primary driver of technological innovation or funder of research and development (R&D).
- While democracies provide greater freedom for scientific exploration, authoritarian regimes can direct sustained funding towards—and maintain focus on—technologies they believe are most important.

The Goldilocks Challenge: Moving Too Quickly, Moving Too Slowly

Technological progress creates risks related to speed. Moving too quickly can disrupt understandings, written or unwritten, that balance a variety of legitimate national, organizational, and personal interests. Rapid or accelerating change could also have a negative impact on safety, security, employment, ethical considerations, societal impacts, and geopolitics. The result could be a public backlash against a particular technology. For example, genetically modified organisms (GMOs) have faced public resistance in Europe due to safety concerns, the Concorde supersonic aircraft was retired over noise and cost issues, and calls for artificial intelligence (AI) regulation reflect fears of the technology's societal impact.

Conversely, innovation that is too slow increases the likelihood that a nation will lose the technical, economic, and national security advantages that often accrue to first movers in a field. Such concerns are apparent in reports asserting that the United States is falling behind China in the development of key technologies considered critical to both national security and economic security, such as AI.[1]

To fully realize the benefits of innovation, policy measures must address both the risks of rapid change and the dangers of falling behind.

Increasing Access to New Technologies Worldwide

A fundamental reality of today's technological environment is that American-born innovations are unlikely to remain in the exclusive control of American actors for long. The diffusion of many of these technologies is, in part, driven by the long-term trend of decreasing information technology costs, but other factors play important roles as well.

Access to and use of these technologies has spread beyond US borders because of global business models that have increased the potential customer base by leaps and bounds.[2] Digital platforms and strong network effects have driven rapid, global user adoption.[3] Open-source initiatives and collaborative research have accelerated diffusion of the underlying technologies by lowering entry barriers and encouraging adaptation across borders.[4] Offshore manufacturing of American-designed innovations and licensing of these innovations has brought technical know-how within the reach of potential overseas competitors.[5] Foreign competitors steal US intellectual property worth hundreds of billions of dollars per year.[6] Technological knowledge is often reverse-engineered or reimagined internationally.[7]

Several key implications arise from increasing access to new technologies:

- Winning isn't winning anymore. The old model of achieving lasting national technological dominance is being replaced by a paradigm of continuous competition where technological advantages are rarely, if ever, sustained for long periods.
- More state and nonstate actors are obtaining access to advanced technologies, gaining new tools to challenge US interests. This makes formulating policies even more complex.
- Technological advantages are narrowing, even on the frontier. Although the United States may possess the most technologically advanced capabilities in certain domains, other actors with less sophisticated—but still effective—versions of advanced technologies can reduce first-mover advantages the US previously enjoyed.
- There are more actors with different ethical thresholds, constraints, and perspectives. Those with fewer bureaucratic and ethical constraints may exploit and adapt technology faster and more effectively than those with more stringent regulations.

To be sure, there are exceptions to this trend of faster and wider technological diffusion. Perhaps the most important of them are instances when scale is a critical aspect of widespread innovation and those in which actors lack access to the natural resources (such as rare-earth metals) or financial capital needed to support large-scale deployments. This has been true for much of the past with respect to nuclear weapons, where the major roadblock for nations seeking to acquire such weapons has been getting access to fissile materials rather than to necessary knowledge. It is also true in AI today, where a small number of private-sector actors clearly dominate the creation of large language models (LLMs).

It may be possible to extend periods of American monopoly on certain technologies (e.g., through the application of export controls to key components of them), but these periods cannot be prolonged indefinitely. Extensions can help to buy time for US policymakers to better anticipate the consequences of a technology's diffusion in the future. But all too often, buying time becomes an end unto itself, and actions to craft a better policy that could help sustain US leadership in key domains—such as targeted immigration reform to attract more of the world's best talent and create a "brain gain" for American universities and companies—are not taken.

The Changing Role of Government in Technological Innovation

Many technological innovations, including satellites, jet engines, and semiconductors, have their roots in US government financial support and advocacy. But in many fields today, the US government is no longer the primary driver of innovation. Private companies have taken up much of the slack. These businesses, however, may be under the jurisdiction of nations—or controlled by senior executives—whose interests are not aligned with those of the users of their services. For example, the Starlink satellite communications network has been an essential part of Ukrainian battlefield communications; however, the CEO of Starlink has curtailed Ukrainian access on a number of occasions in ways that affected Ukraine's battlefield strategy.[8] Such concerns are most serious when there is only one—or just a small number—of private-sector providers of the services in question.

No better example of private companies' growing influence in setting the R&D agenda can be found than in the current scene for funding AI research. Whereas the federal government talks in terms of billions of dollars in federal support for AI research, the private sector talks in terms of amounts a hundred or more times larger. Similar trends seem to apply to biotechnology and synthetic biology research, though not quite as starkly. And, as chapter 10, on space, discusses, services related to space are increasingly being delivered by private companies.

EXTREME ULTRAVIOLET AND ADVANCED CHIP FABRICATION

Advanced chip fabrication is a domain of national technology policy in which developments appropriate thirty years ago may need reassessment today. The most advanced chips currently being made require light in the extreme ultraviolet (EUV) range (13.5 nanometers). The method used today for producing EUV light uses high-energy laser pulses to vaporize tiny droplets of tin to create a plasma that emits EUV light. This light is then precisely reflected and focused by some of the flattest mirrors in the world to etch the intricate patterns needed for advanced semiconductor chips. This technology is key to increasing circuit density on advanced chips, making them faster and more powerful.

Major breakthroughs in this laser technology were developed by researchers at Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and Sandia National Laboratory in the 1990s, and intellectual property rights were owned by the US government but licensed under approval by Congress and the US Department of Energy. The Dutch company ASML applied for a license, and at the time no objections were raised.

Today, ASML is the only company in the world that can manufacture and service the sophisticated machines using EUV technology. Regular EUV machines cost about $200 million each, but the newer high-numerical-aperture EUV systems cost closer to $370 million, and ASML can manufacture only a handful annually.[a] The future development of advanced semiconductor manufacturing equipment that will increase circuit density on chips even more—as well as the balancing of market access with the national security concerns of exporting to and servicing ASML equipment in China—are major geopolitical and economic concerns.[b]

a. Mat Honan and James O'Donnell, "How ASML Took Over the Chipmaking Chessboard," *MIT Technology Review*, April 1, 2024, https://www.technologyreview.com/2024/04/01/1090393/how-asml-took-over-the-chipmaking-chessboard/; Charlotte Trueman, "Intel Acquires ASML's Entire 2024 Stock of High NA EUV Machines," *Data Center Dynamics*, May 9, 2024, https://www.datacenterdynamics.com/en/news/intel-acquires-asmls-entire-2024-stock-of-high-na-euv-machines/.

b. Arjun Kharpal, "Netherlands Takes On U.S. Export Controls, Controlling Shipments of Some ASML Machines," CNBC, September 6, 2024, https://www.cnbc.com/2024/09/06/netherlands-expands-export-curbs-on-advanced-chip-tools.html.

The growing influence of the private sector in critical technologies has led US officials to emphasize the need for closer public-private cooperation and government regulation. Even if the government does not lead in innovation, it still plays a crucial role in funding R&D—and especially R&D with lengthy time horizons—promoting key innovations, setting standards, and stimulating the formation of coalitions of private-sector actors domestically and internationally.

The Relationship of Political Regime Type to Technological Progress

National priorities can change with the evolution of the geopolitical environment. In the 1990s, there was widespread optimism about the triumph of liberal democracy and free market capitalism. Much of US economic policy was characterized by efforts to support free trade, accelerate globalization, and promote China's integration into the world economy as a way of facilitating the country's transition to more democratic rule.

During this time, the global manufacturing landscape for key technologies, particularly semiconductors, underwent significant shifts. Over the past three decades, the US share of global semiconductor production dropped from 37 to 12 percent, as noted in chapter 9, on semiconductors. Meanwhile, Asian manufacturers, especially in South Korea and

Taiwan, have emerged as major players, supported by government policies and regional demand shifts. Asia has become the dominant region for semiconductor production, laying the groundwork for the current global supply chain.

This shift in manufacturing capabilities, coupled with China's economic and military rise, is a key element of changes in the geopolitical environment, and it drives many Western concerns about technological dependencies in the twenty-first century. Accordingly, national policies that were seen as useful and appropriate in the environment of thirty years ago may need reassessment today. (See the sidebar on extreme ultraviolet and advanced chip fabrication for an example.)

Finally, although genuine technological innovation occurs in both democracies and autocracies, each regime type has different advantages and faces different challenges. Democracies benefit from the rule of law, a free flow of ideas and people, and greater freedom for individuals to pursue their own research interests. Perhaps most importantly, because failure in a democracy should not lead to persecution or professional ostracism, individuals are freer to experiment and explore. By contrast, authoritarian regimes are characterized by the rule of the state and sometimes impose dire consequences for failure, which can restrict the flow of ideas, force adherence to state-approved research areas, and prompt scientists to focus only on what a government considers safe topics.

On the other hand, authoritarian regimes can direct sustained funding and attention to areas deemed crucial by the state more easily than democracies can; they can also maintain focus on these areas for extended periods, independent of short-term profit or political considerations. For example, it is widely accepted that Chinese AI efforts have access to the personal data of individuals on a far broader scale than such efforts in the West, which generally has stronger privacy protections against governmental intrusion than China does.

Innovation Pathways and Patterns of Progress

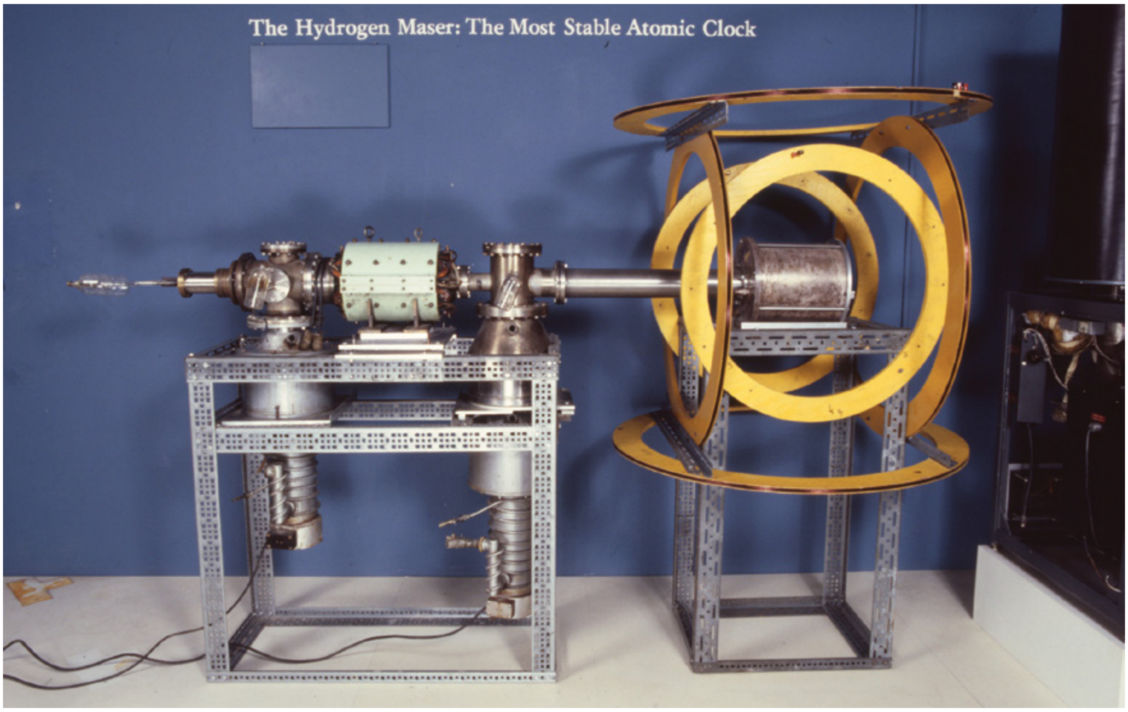
KEY TAKEAWAYS

- Technological progress is often unpredictable and nonlinear, with periods of slow development interrupted by sudden breakthroughs. While some fields, like semiconductors, have shown steady improvement, most technologies advance through cycles of experimentation, feedback, and convergence of multiple innovations.
- Nonscientific factors, such as engineering feasibility, economic viability, manufacturing challenges, and societal acceptance, influence the adoption of technology based on scientific advances.
- Hype can distort perceptions, leading to inflated expectations that outpace practical utility and distortions in resource allocation.
- Frontier bias causes overemphasis on new technologies and sometimes results in overlooking impactful uses of established ones.
- The synergies between different technologies are large and growing, which makes understanding the interactions between different fields all the more important.

The Unpredictable and Nonlinear Nature of Technological Progress

Technological progress exhibits a variety of patterns. For example, progress in semiconductors has been fairly predictable historically, progressing consistently with Moore's law, which predicts a continuing exponential decrease in the cost of computation over time. But, as noted in chapter 9 on semiconductors, this steady decline is coming to an end, if it hasn't expired already. Solar cells and light-emitting diode (LED) lighting have followed similar cost-reduction

FIGURE 11.1 Norman Ramsey's hydrogen maser, made in 1959–60, known as the "most stable atomic clock," displayed at the Smithsonian's National Museum of American History



Source: The Smithsonian's National Museum of American History, https://americanhistory.si.edu/collections/object/nmah_714239

curves, except that these cost decreases are usually represented as a function of manufacturing experience and expertise rather than of time.[9]

Most other technologies have demonstrated much more uneven progress, characterized by extended phases of gradual development interrupted by sudden, transformative bursts of innovation. Sometimes, these bursts result from particular breakthroughs. For example:

- Timekeeping relied on devices such as sundials and water clocks until the invention of the mechanical clock in the late Middle Ages provided the first significant jump in accuracy.[10] Following a prolonged period of incremental refinement, the introduction of electric clocks—with the frequency of alternating current (AC) from the wall providing a time base—further enhanced the accuracy and reliability of clocks. Subsequent innovations, including quartz and atomic clocks (see figure 11.1), established today's remarkable standards of precision.
- Crop yields have increased only incrementally over time, except during certain periods of rapid innovation associated with technological developments such as synthetic fertilizers, mechanization, high-yield crop varieties, and the rise of biotechnology and precision agriculture.[11] During these periods, crop yields jumped quite substantially, only to resume a pace of gradual improvements as these innovations spread more widely.

- The World Wide Web emerged in the 1990s as a significant development in global communication and information exchange.[12] The web enabled users to access and navigate information through interconnected hypertext links. Used with point-and-click interfaces (also known as browsers), it was rapidly adopted. This led to substantial growth in websites, online communities, and e-commerce and influenced worldwide information accessibility and interaction.

At other times, a surge in innovation is due to the simultaneous availability and maturity of several key technologies that are combined to achieve significant progress in some other technological domain—this is the convergence phenomenon, discussed later in this section.

Predicting future progress can be challenging and misleading due to this pattern of punctuated innovation. Even experts in a given field can be surprised by the rapidity of progress. For instance, Geoffrey Hinton, a pioneer in AI and a winner of the 2024 Nobel Prize in Physics for his application of tools and concepts from statistical mechanics to machine learning, recently expressed astonishment at the swift progress in AI and predicted that it will surpass human intelligence in the future.

His comments came after a long history of multiple "AI winters."[13] Enthusiasm for AI in the 1950s and 1960s subsequently led to the first major AI winter (1974–80). The 1980s saw a revival of enthusiasm for AI involving rule-based expert systems, but unmet expectations triggered a second winter in the late 1980s and into the 1990s. Progress inspired by the machine learning approach in the 2000s led to a resurgence in the next two decades. This led to the new surge of enthusiasm and optimism we are witnessing today, which is driven by advances in deep learning, very large datasets, and increases in computing power.

The punctuated nature of most technological change suggests that expectations of regular and rapid evolution in many fields are generally not realized,[14] despite what headlines in the news might lead one to believe. This point is also relevant to another important observation: The traditional linear model of R&D, which envisages smooth progress from basic research to applied research, leading to development and then to marketable products, represents just one way in which societies derive value from technological investments.

Progress also occurs in nonlinear ways that depend on feedback between the various stages of activity. For example, some challenging problems require a deeper fundamental scientific understanding known as "use-inspired basic research," which comes into play after innovations have already been deployed. Research in AI on LLM hallucinations (outputs of entirely false statements) fits into this category. The models are already broadly useful despite such errors occurring frequently. Nevertheless, these hallucinations are problematic, and important research is underway to understand the mechanisms that lead LLMs to generate them.

In other cases, technology convergence can have a big impact on synergy and innovation. Here, convergence means that several distinct technologies have advanced to the point at which they can be integrated to develop a useful innovation. For example, electric cars today are made possible by the convergence of advances in battery technology, lightweight materials, sensors, and computing power. Together, these advances have improved vehicle range, safety, and efficiency and have enabled features like autonomous driving and real-time diagnostics. Another example comes from chapter 6 in this report, on neuroscience, which discusses how effective neurological interventions depend not only on a fundamental theoretical understanding of brain function, but also on the development of neural probes that can be implanted into the brain without causing serious damage to brain tissue.

In short, for most applications, true innovation requires repeated cycles of experimentation, learning, and adaptation rather than a single, direct path. Feedback, convergence, and iteration are the norm, not the exception.

Nonscientific Influences on Innovation

Scientific advances are frequently highlighted for their promise to address societal challenges and enhance our quality of life. However, there is often a large gap between a demonstration of scientific feasibility and the creation of an economically viable and societally useful product or service based on the technology.

After achieving scientific proof of concept, a given technology application based on that science must demonstrate engineering practicality. An example is the idea of a chemically fueled, single-stage-to-orbit spacecraft launched from Earth. It is generally believed that launching a spacecraft using a single rocket stage, rather than multiple stages, is just barely possible using current rocket fuels and materials. However, not even a leading company like SpaceX has been able to demonstrate a feasible engineering design that could reliably accomplish this task.

Economic viability and practicality come after engineering feasibility, and these involve considerations such as cost and ease of use. Early attempts to build supercomputers with superconducting components demonstrated technical success but faced practical challenges due to the need for liquid helium for cooling. This requirement made the computers difficult and costly to deploy, and the development of alternative technologies offering comparable performance at lower cost doomed the approach in the marketplace.

Manufacturing comes next. Even if engineering feasibility has been demonstrated, developing a viable manufacturing process to build a product or service based on the initial scientific proof of concept may still prove too difficult. There may be other constraints as well: For instance, materials used to demonstrate engineering feasibility may be too expensive or rare to support large-scale production. (Manufacturing is discussed in more detail later in this chapter.)

Another important factor is the availability of cheaper alternatives to a new technology, which can undermine the commercial viability of the innovation. The competition between lithium-ion (Li-ion) and sodium-ion batteries illustrates this. Sodium-ion batteries have potential cost advantages over Li-ion ones because sodium is much more abundant than lithium. But Li-ion battery technology has a head start of a couple of decades, and work on producing these batteries has driven down their cost signficantly. Thus, the economics of procurement today favor Li-ion batteries in many common applications. However, any significant disruption of the lithium supply chain could make sodium-ion ones more competitive.

Societal acceptability is yet another important nonscientific influence on technological progress. The psychology of individuals, as well as the cultural practices and beliefs of a community or society, influence the adoption and use of any given application of an emerging technology. For instance, producing and consuming GMOs as food is highly controversial in Europe, and concerns over their safety have prevented the uptake there of GMO foods that are consumed widely in the United States.

Finally, the journey from scientific breakthrough to practical, widespread application is often more difficult than anticipated. Innovators may discover that fulfilling promises to investors and customers requires greater resources and longer timelines and delivers fewer benefits or capabilities than expected. Obstacles such as raising adequate funding, navigating environmental or social concerns, and managing risks related to ethics, privacy, and public trust frequently surface only as products or services reach the market. Policymakers, therefore, face the complex challenge of supporting promising advances while being mindful of their associated risks.

Striking the right balance requires acknowledging that disruptive technological progress brings both opportunity and uncertainty. While advocates may downplay early concerns as barriers to the innovations that would benefit their businesses, unaddressed risks—especially ones that could impact society—can escalate as new technologies scale. Governments can mitigate these challenges by incorporating diverse and even critical perspectives early in the technology life cycle, fostering an environment that encourages innovation while managing its potential downsides.

Technological Optimism, Hyperbole, and Technical Reality

This publication highlights ten significant emerging technologies. In preparing the latest edition of the *Stanford Emerging Technology Review* (*SETR*), faculty members from each of these fields expressed broad optimism about the societal and scientific value of research in their respective domains. This optimism is hardly surprising: Those who dedicate their careers to advancing new technologies naturally believe in their potential to address important challenges. Indeed, a conviction that progress will continue and yield solutions is almost a prerequisite for anyone deeply invested in innovation.

However, the line between responsible optimism and irresponsible hype can be crossed easily, leading to an all-too-common pattern where media coverage ignoring basic scientific fundamentals leads to overinflated expectations among the public.

Technological hype often begins with a breakthrough in an emerging technology area, quickly followed by grandiose promises of disruptive, even revolutionary impact. Such promises—in reality, overpromises—make claims that go far beyond what available knowledge and evidence support. They focus on potential rather than proven functionality and often rely on emotional appeal and ambiguous terminology. They also imply that all social or economic challenges can be solved through technological innovation alone, ignoring implementation barriers, regulatory considerations, social acceptance, or the harms often caused by large-scale deployments of unproven technology.

A prominent example of technological hyperbole comes from 1989, when two chemists at the University of Utah announced they had achieved cold fusion—that is, fusion reactions at low temperatures. This finding, if true, would have challenged the scientific consensus that extremely high temperatures are required for such reactions. Rather than following the standard scientific process of peer review, the researchers revealed their findings at a press conference, touting the potential for a clean, virtually inexhaustible energy source.[15] National advisors emphasized the discovery's importance, suggesting that it was too significant to leave solely to the scientific community.[16] The implication was that factors other than science should play an important role in how the nation should proceed at that moment. This was despite the fact that the possibility being discussed (that cold fusion had actually been discovered) was entirely a scientific question.

The scientists' subsequent publication underwent expedited peer review and was widely criticized for lacking essential experimental details, which made independent verification difficult. Their claims rested on observations of heat production that they attributed to fusion. However, later investigations identified significant flaws in their measurement techniques, undermining confidence in their overall findings.

This 1989 episode is widely regarded today as an object lesson in the perils of circumventing the normal processes of science. Some researchers are still working on low-temperature fusion as a plausible mechanism for generating energy, and the field is supported at the level of around $10 million per year in research funding. However, this level of support—a very small fraction of the funding dedicated to more traditional fusion research—should not be regarded as vindication for the original cold fusion proponents. Rather, it reflects the quite modest level of support appropriate for an approach to fusion that is regarded skeptically by most in the scientific community but has not been shown to be categorically false.

Technological hype affects investors, consumers, policymakers, and other stakeholders—all of whom must navigate the complex interplay between marketing rhetoric and substantive advancement. Underlying any given instance of technological hype is often something genuine—some scientific or technological development that is in fact new or noteworthy. But public stakeholders would be well advised to allow the scientific review process to play out before jumping on a hyperbole-driven bandwagon.

Frontier Bias

Frontier bias is a tendency among analysts, commentators, and policymakers to focus on the significance of the newest and most recent innovations. Such a trend has been apparent even in the uptake of earlier versions of this report—requests for briefings arising from the publications have most often focused on what's newest and most advanced in various fields. Frontier bias emerges from many sources, but one of the most prominent is the technology hype described in the previous section.

Given humans' predilection for novelty, this bias is understandable. But it carries with it the risk of overlooking "old" technologies that can be used in novel and impactful ways. Innovation using proven and known technologies is a powerful means of advancing national and societal interests and, by definition, does not rely on fundamental scientific or technological breakthroughs.

One prominent example of older or known technologies being used in such ways can be seen in the present Russia-Ukraine war. Many of the drones having a significant effect on the battlefield are a diverse mix of moderately sophisticated ones and off-the-shelf commercial drones. And, in response to US trade sanctions on advanced semiconductors, Russia is using chips designed for home and commercial use to control its weapons.

Another example is the widespread use of the AK-47 automatic rifle. Unlike other popular guns, the AK-47 was deliberately designed to be low-tech—cheap, simple, and durable; easy to manufacture; and with few moving parts. It has since proliferated: Some seventy-five million of these guns are in operation today, and they have been widely adopted by forces around the world,[17] most notably insurgent groups and terrorists.[18]

The story in chapter 10 on sustainable energy in *SETR* 2025, about a second life for electric vehicle (EV) batteries, is also relevant. As EVs become more prevalent, batteries in them that are coming to the end of their useful life face being discarded. But they often still have significant capacity for power storage. Specialized battery-management systems tailored to these batteries' unique characteristics can help them serve in stationary energy-storage applications (see figure 11.2), such as by acting as

[Frontier] bias is understandable. But it carries with it the risk of overlooking "old" technologies that can be used in novel and impactful ways.

FIGURE 11.2 Advanced battery-management systems being used as stationary storage capacity

Source: Smartville

backup power sources for the grid. Because of their age, they may not be at the cutting edge of battery technology when they are converted, but such systems can give a productive second life to batteries that would otherwise be thrown away.

A second consequence of frontier bias is a misunderstanding of the difference between scientific or technological advances and adoption at scale—a phenomenon that was noted earlier under "nonscientific influences on innovation." For example, in the couple of decades after the first generation of commercial nuclear power in October 1956, there was considerable optimism that further technological advancements in the field would bring about an era in which electrical energy was too cheap to meter. But, as discussed in chapter 10 in *SETR* 2025, nuclear fission has not been widely adopted as a source of energy for a variety of technical, economic, and political reasons.

For an innovation to have significant societal impact, it needs to be broadly available and widely used. At one extreme, some innovations can be acquired by individuals based on their own personal needs. The rapid spread of personal computers in the 1980s and of rooftop solar panels for home electricity generation in the past decade are examples of people willing to spend money out of their own pockets to derive the benefits of these innovations. The result was rapid uptake and adoption throughout society.

At the other extreme, advanced technology that needs a significant degree of centralized planning or funding for realization is likely to require much longer timescales for widespread adoption. Nuclear energy requires the construction of nuclear reactors that cost billions of dollars. State-of-the-art semiconductor plants cost tens of billions of dollars. Medicines for treating neurodegeneration are available only at the end of a very expensive drug-approval and manufacturing process. Carbon capture and sequestration is too expensive to be widely adopted and is of marginal benefit to individuals, though it is of use to industrial facilities. For such innovations, it is unrealistic to expect rapid and widespread adoption throughout society.

Large and Growing Synergies Between Different Technologies

The synergies between different technologies are both significant and expanding, as advances in one field often enhance progress in others. For example:

- Artificial intelligence (AI) contributes to advances in synthetic biology by predicting the structures of various biomolecules, such as proteins, nucleic acids, and small molecules singly or joined together in various complexes.[19]
- AI helps to screen many candidate compounds to predict the ones most likely to exhibit desirable properties for materials science.[20]
- Materials science is central to the identification of new semiconductors that may be useful in developing more energy-efficient chips, which in turn can reduce the cost of training AI models.[21]
- Materials science is important in space research, where the creation of new materials may be needed for the construction of advanced spacecraft and satellites.[22] It is also important in neuroscience, where it enables the development of neural probes that can send and receive electrical signals in neural tissue.[23]
- Energy technologies help to improve the performance of robotics and spacecraft.[24]
- Synthetic biology can build organisms that produce certain specialized materials.[25]
- Cheaper semiconductors have driven down the cost of DNA sequencing, which itself is a fundamental technology for synthetic biology.[26]

This point is more obvious when a field such as AI or materials science is seen as a technology that impacts a variety of application domains. For example, this report has discussed how AI has facilitated innovations in battery technology and in protein folding. Less obvious is that AI itself has benefited greatly from advances in semiconductor technology, which has itself benefited from developments in materials science.

In certain instances, a useful technology becomes an enabling technology—a technology whose existence and characteristics enable applications that would not otherwise be feasible or affordable, especially across a number of different fields.[27] (The sidebar on lasers as an enabling technology across multiple sectors, drawn from *SETR* 2025, provides an example.)

An enabling technology can evolve into a general-purpose technology if it becomes broadly useful across many domains. A general-purpose technology is characterized by continuing improvement, wide applicability, and benefits that extend well beyond its original uses. Each advance in a general-purpose technology amplifies its overall impact. Historical examples—such as the steam engine, electricity, and information technology—have transformed economic growth, industry, and daily life. General-purpose technologies ultimately reshape how people, firms, and governments interact with a wide range of other technologies and with one another.

Human Capital and Knowledge Ecosystems

KEY TAKEAWAYS

- Human capital is the foundation of scientific and technological progress. Sustained investment in it is the single most critical factor in ensuring long-term national competitiveness and scientific advancement.
- Universities are central both to high-risk research and to science, technology, engineering, and mathematics (STEM) education. Yet federal R&D funding as a share of GDP has declined, and policy ambiguities hinder international collaboration.

LASERS: AN ENABLING TECHNOLOGY ACROSS MULTIPLE SECTORS

Lasers, as highlighted in the 2025 edition of the *Stanford Emerging Technology Review,* are an enabling technology for a wide array of scientific and industrial fields due to their precision, versatility, and efficiency.

Medicine

- **Surgical precision** Lasers are used to ablate, cut, or vaporize tissue and to clot bodily fluids. Unlike traditional tools such as saws or drills, lasers provide cleaner, more precise cuts, minimizing mechanical and thermal damage to surrounding tissues.
- **Cancer treatment** Lasers can target and destroy subsurface tumors with minimal harm to healthy tissue, offering less invasive alternatives for certain procedures.

Military applications

- **Directed-energy weapons** Lasers are being developed as weapons capable of disabling satellites and providing short-range air defense against drones, rockets, and artillery.
- **Target designation** Lasers play a crucial role in guiding munitions by marking targets with beams of light, allowing for highly accurate strikes.

Communications

- **Fiber-optic data transmission** Lasers transmit vast amounts of data through fiber-optic cables. Advances now allow for much shorter laser pulses, maintaining data fidelity while potentially reducing power consumption.
- **Satellite links** Lasers enable high-speed, long-range data transmission between satellites, supporting global communications infrastructure.

Manufacturing

- **3-D printing** Lasers are integral to additive manufacturing techniques such as stereolithography and selective laser sintering. In stereolithography, ultraviolet lasers cure photosensitive resin layer by layer, while in selective laser sintering, lasers fuse powdered materials like nylon or metal. These methods allow for rapid prototyping and the creation of complex structures from various materials.

Imaging

- **X-ray free-electron lasers (XFELs)** XFELs generate powerful X-ray pulses that penetrate materials, enabling high-resolution imaging and measurement of physical properties. Their short wavelengths provide superior spatial resolution compared to visible light, facilitating breakthroughs such as imaging new proteins, observing quantum material phase transitions, and tracking biomolecular movements in real time.

- The "valley of death" between research feasibility and commercial viability remains a major barrier to advancing innovations to market. New funding models are needed to bridge this gap and sustain America's technological leadership.

The Central Importance of Ideas and Human Talent in Science and Technology

Scientific progress thrives on new ideas, which are generated daily by the most talented individuals worldwide. But human talent capable of creating ideas in science and technology cannot be generated on demand. Such talent must be nurtured domestically or acquired from foreign sources.

Workers of US origin still make up the majority of the US STEM workforce, although foreign-born talent accounts for an increasingly large fraction of it. Strengthening the domestic pipeline of STEM workers is essential for several reasons.

- First, a number of studies indicate a strong correlation between a nation's STEM education and economic growth and productivity.[28] Correlation is not causation, but the connection is unlikely to be accidental or spurious.

- Second, other nations—such as China and India, from which significant numbers of STEM students in the United States originate—are investing more heavily in scientific R&D. Individuals who have previously chosen to work and study in the United States may well take advantage of opportunities at home created by such investment in greater numbers. Foreign-born individuals working in the US STEM workforce may have family or personal ties in their nations of origin that tempt them to return. Foreign countries may also take steps that explicitly discourage their scientists and engineers from studying or working in the United States.

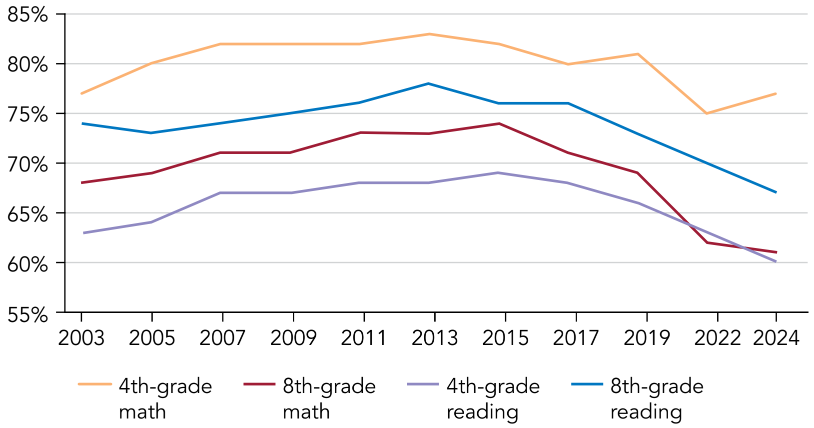
- Third, many security-sensitive jobs depend on US citizens. In 2021, the US Department of Defense (DOD) noted that improving the capacity and resilience of the defense industrial base requires more workers trained in STEM.[29] It also observed that the dearth of trained software engineers working on classified projects was in part because of the requirement that they are US citizens. In 2025, the aerospace and defense sector continued to face a severe talent shortfall, with industry analysts estimating that about fifty thousand software and technology positions remain unfilled.[30]

According to analysts from the National Defense Industrial Association's Emerging Technologies Institute and the Institute for Progress,[31] the US defense industrial base relies on roughly 110,000 foreign-born STEM graduates at any given time; of this number, 85 percent are naturalized citizens. As they conclude, "[US] Defense Department projects are disproportionately likely to turn to international talent [i.e., talent from foreign sources] for advanced STEM skills."

In promoting a more robust domestic contribution to building STEM expertise in America, it is sobering to realize that the United States is also facing a decades-long decline in K–12 (kindergarten to twelfth grade) STEM proficiency,[32] with standardized testing revealing declining scores in fourth- and eighth-grade mathematics.[33] While COVID-19 disruptions account for some of the decline,[34] the 2024 National Assessment of Educational Progress (released in January 2025) shows that US math scores remain below pre-pandemic levels (as seen in figure 11.3): Fourth-grade math scores have risen since 2022 but were still below their 2019 level, while eighth-grade math scores dropped compared with 2019. Reading scores fell from 2019 levels for both grades.[35] This follows a twenty-year trend of diminishing US K–12 STEM proficiency.[36]

Another data point is found in the five-year trend from the national ACT (American College Testing) test, a curriculum-based assessment of high school seniors tracking the mastery of college-readiness standards.[37]

FIGURE 11.3 National Assessment of Educational Progress scores over time



Source: National Assessment of Educational Progress

Since 2020, scores on the college-readiness benchmarks for mathematics and science have dropped monotonically. In 2024, only 29 percent of seniors met the readiness standard for mathematics, and only 30 percent met the standard for science, highlighting a critical educational challenge for the nation's economic and technological competitiveness.

Of particular concern is that only 7 percent of American teens scored in the highest level of math proficiency as measured in 2022 by the Program for International Student Assessment, a test to assess student ability to apply knowledge in real-world situations, administered by the Organisation for Economic Co-operation and Development. This is compared to 12 percent of Canadians and 41 percent of Singaporean teens scoring in the top category.[38]

Adding to the challenge is a shortage of qualified STEM educators in the United States. Even as early as 2012, about 30 percent of math teachers, 26 percent of biology teachers, and 54 percent of physical science teachers lacked a major or degree relevant to their teaching assignment,[39] and there is no reason to believe that the situation has improved since then. Further, one study from 2021 estimates the shortage of qualified STEM teachers in middle and high school at between 180,000 and 350,000.[40] Simultaneously, the annual production of new STEM teachers in America has declined, falling from about 31,000 a decade ago to roughly 20,000 today.[41]

When considering foreign sources of STEM talent, immigration policies affecting the labor force can make it harder to meet recruitment goals in industries like semiconductors, biotechnology, and sustainable energy. Foreign talent makes critical contributions to US STEM.

R&D funding levels are also changing. Although America remains the single most prominent contributor to global R&D, other nations—most notably China—are rapidly increasing their investments in this area. Geographic concentration of R&D expenditure continues its shift from the United States and Europe to East, Southeast, and South Asia.

This trend highlights the increasing importance of international collaboration. US researchers benefit from ideas developed abroad when they read scientific literature from other countries, but direct interactions with foreign researchers are often more valuable because they provide more comprehensive and expansive insights. Such interactions help American researchers acquire tacit knowledge that is not captured in published papers, including research directions that appeared promising but did not ultimately bear fruit. They also offer a deeper understanding of foreign scientific progress. (See the sidebar on the importance of tacit knowledge for more information.)

This point about the importance of tacit knowledge in scientific advancement has been made by many scholars,[42] and it was expressed particularly strongly in a multitude of interviews with Stanford faculty working in the technology areas addressed in this report.

America's ability to attract and retain foreign talent is essential for maintaining its innovation edge, and domestic innovation is hindered when limitations are imposed on interactions with foreign scientists and their research. Skilled immigrants play a crucial role in American innovation, with immigrant college graduates receiving patents at twice the rate of native-born Americans.[43] More generally, drawing from a broader pool of people will yield higher quality talent than digging more deeply into an existing pool simply because the broader pool is more likely to have a greater number of individuals at the high end of the talent distribution.

THE IMPORTANCE OF TACIT KNOWLEDGE

Tacit knowledge is almost always found at the frontiers of new technologies. Unlike explicit knowledge, which can be codified and shared in documents, tacit knowledge consists of personal, intuitive skills and insights gained through experience and practice that are hard to articulate. Such knowledge allows practitioners to interpret results, troubleshoot equipment, and apply theories effectively—skills that cannot be fully acquired from published papers alone. Instead, these abilities often require working alongside experienced professionals and absorbing problem-solving habits through direct interaction.

The significance of tacit knowledge is clear even outside the laboratory. For example, in the semiconductor sector, close on-site collaboration between chip buyers and manufacturers helps minimize production downtime since diagnosing problems often depends on hands-on familiarity with complex equipment. Some semiconductor companies even embed technicians with their customers worldwide because the technicians' subtle skills at equipment calibration cannot be captured in manuals; rather, they must be conveyed through mentorship and direct, hands-on involvement.

As a technology matures, the tacit knowledge of experts in the field becomes more explicit. This shift signals progress: For any technology to be integrated into a society's infrastructure, informal know-how must be documented and standardized. Turning these unspoken practices into clear procedures and guidance supports wider adoption and also ensures that a technology can be taught, replicated, and relied upon by a broader group of practitioners. This transformation—from personal mastery to public instruction—marks the transition from a niche innovation to a stable, essential technology.

US policies that discourage immigration can reduce the influx of skilled workers, impacting the country's capacity for innovation.[44] They also shift skilled talent and multinational R&D investment to other countries, including strategic competitors such as China and also close allies such as Canada.[45] This shift can sometimes force US companies to relocate abroad due to worker shortages.[46]

Finally, many academic researchers are immigrants on student visas. STEM workers educated in the United States are more likely to have personal and, in many cases, citizenship loyalties to America and can fairly be regarded as more likely to remain in the country than to leave after completing their studies. But without offering them a clear route to permanent residence, the United States loses key teaching and research talent in vital STEM domains.

Today, both domestic and foreign paths to growing the requisite talent base to sustain and grow US innovation face serious and rising challenges. The global competition for talent means that the United States must adopt a more strategic approach to leveraging international expertise, as connections between American science and technology efforts and those of the rest of the world will accelerate the nation's progress in critical technology fields. To maintain and enhance its innovation capacity, the United States urgently needs to improve its own STEM education across all demographic groups, provide better pathways for skilled immigrants to remain in the United States, and invest more in human capital.

Concerns about foreign appropriation of American intellectual efforts are not without foundation. But using a meat axe to make blunt, widespread cuts in opportunities for collaboration with foreign scientists when a surgical scalpel could be used to address only the issues warranting serious concern is a sure way to undermine the effectiveness of US scientific endeavors.

Role of Universities in Technological Innovation

Within the innovation ecosystem, universities play two unique and pivotal roles that are often underappreciated. First, they have the mission of pursuing high-risk research openly that may not pay off in commercial or societal applications for a long time, if ever.[47] (See the sidebar on the long-term reach of university research for some examples.) This openness accelerates discovery by making study details, data, and results accessible to others. Private companies contribute to the innovation process, but universities and other research institutions are key to many advancements. One significant data point is that more than 80 percent of the algorithms used today—not just in AI but in all kinds of information technology—originated from sources other than industrial research.[48] University openness magnifies educational and societal benefits by enabling other researchers to build on prior work, thus driving innovation forward.

Second, as educational institutions, universities play the central role in developing STEM expertise within the next generation. Any long-term plan for STEM leadership globally must include efforts to sustain advantages that the United States has. For example, US higher education in STEM is still the best in the world. This leadership is reinforced by the strength of America's university-based research enterprise: There is no better way to learn how to do state-of-the-art research in STEM than to actively participate in such work. By providing students with hands-on research experiences, access to cutting-edge facilities, and mentorship from leading experts, US universities can create an environment where the next generation of STEM leaders will flourish.

Throughout history, government-supported university research has played a key role in technological advancements, from radar and proximity fuses during World War II to modern developments like AI and mRNA vaccines. It has generated knowledge whose exploitation has created new industries and jobs, spurred economic growth, and supported a high standard of living while also achieving national goals for defense, health, and energy.[49] It has also been a rich source of new ideas, particularly for the longer term, and universities are the primary source of graduates with advanced science and technology skills.

University R&D funding from all sources grew significantly in 2023, reaching $108.8 billion—an 11.2 percent increase from 2022[50] (though it is likely to now be significantly lower given recent funding cuts). While private-sector investment in technology and university research has increased, it cannot replace federal funding, which supports R&D focused on national and public issues rather than on commercial viability.[51] The US government remains uniquely capable of making large investments year after year in basic science at universities and national laboratories, which is essential for future applications. Nevertheless, the proportion of academic R&D funding supported by the US government declined over the past decade, standing at 55 percent of total support for academic R&D in 2023, the most recent year with available data.[52]

As a percentage of GDP, funding trends have also been negative. The fraction of GDP that goes to R&D could fairly be regarded as a seed corn investment in the future, yet federal R&D funding has fallen from 1.86 percent of GDP in 1964 to just 0.63 percent of GDP in 2022.[53]

Until 2025, a constrained budget environment was the primary driver of these negative funding trends. For example, the Creating Helpful Incentives to Produce Semiconductors (CHIPS) and Science Act of 2022 authorized dramatically increased funding for basic research—about $53 billion—but Congress provided only $39 billion in the corresponding appropriation.[54] The United States still funds more basic research than China, but Chinese investment is rising much more rapidly and will likely overtake that of the United States within a decade.[55]

THE LONG-TERM REACH OF UNIVERSITY RESEARCH

Research in number theory—a branch of pure mathematics—was undertaken for decades before it became foundational to modern cryptography. In the 1960s, academic research on perceptrons sought to develop a computational basis for understanding the activity of the human brain. (A perceptron is the simplest form of neural network; it has one layer of artificial neurons.) Although this line of research was abandoned after a decade or so, it ultimately gave rise to the work on deep learning in artificial intelligence several decades later.

The term *mRNA vaccines* entered the public lexicon in 2021 when COVID-19 vaccines were released.[a] Yet development of these vaccines was built on university research with a thirty-year history.

Magnetic resonance imaging (MRI) was first discovered in university studies in the 1940s, but it took another three decades of research, much of it university based, for the first medical MRI imagers to emerge.

a. Elie Dolgin, "The Tangled History of mRNA Vaccines," *Nature*, October 22, 2021, https://www.nature.com/articles/d41586-021-02483-w.

Moreover, despite their vital contributions, universities face challenges due to the blurring line between fundamental and export-controlled research, which complicates international collaboration in fields such as semiconductors, nanotechnology, AI, and neuroscience. For example, some researchers worry that fundamental research, which should be a less sensitive area, could now be considered export controlled, and they may shy away from foreign collaboration out of an abundance of caution. While well intended, these kinds of expanding restrictions may backfire in the long term, holding back US progress in key technological domains. Restrictions are not the only challenge; policy ambiguity is also harmful because it can discourage or deter collaboration with non-American researchers wishing to contribute to work in the United States.

All of these policy issues, widely recognized among the research community and apparent in interviews with Stanford faculty for this publication, underscore the urgent need for clarification and reform to advance research and promote effective international collaborations.

Finally, it is true that the US R&D landscape is vast, with major contributions from both private industry and the federal government. Historically, private research centers like Bell Labs and IBM's Thomas J. Watson Research Center advanced foundational science. However, most corporate R&D today is focused on applied, proprietary work with limited accessibility. Federal labs and other government-backed research facilities, such as those run by the US Department of Energy (DOE), DOD, and NASA, tackle complex, mission-specific challenges. But the research undertaken in the private sector and federal laboratories does not substitute for university research: Unlike mission-driven federal labs, universities pursue a broad range of research topics, and unlike the private sector, they emphasize open, transparent research that fosters accountability, collaboration, and wide-reaching impact.

The Structure of Research and Development Funding

The scale of investment that nations make in R&D matters, but it is also critical how that money is allocated. First, the government plays an important role in funding long-term precompetitive research that industry is not structured to support. Second, frequent shifts in funding levels, which are becoming increasingly common in government funding, undermine systematic R&D efforts and drive away scientific talent that opts to find employment elsewhere. Third, the so-called valley of death, a period after the engineering feasibility of an innovation has been demonstrated but before large-scale adoption and commercial viability has been achieved, is a significant problem.

That valley exists because when a new innovation is first offered to customers, its cost relative to what it is capable of can be a deterrent to adoption. High initial costs can put off the public from purchasing or using the innovation, potentially leading to a firm's commercial failure in the absence of external funding. However, as production volume increases, per-unit costs typically decrease due to the learning curve in manufacturing. This cost reduction is critical, especially in sectors like energy production, where large-scale deployment offers significant societal benefits.

The problem is that researchers and young companies trying to reach this point must first find ways to scale their activities to demonstrate their innovations' capabilities at scale—and raising money to do this can be challenging. Research funding typically ceases once the feasibility of a technology has been demonstrated. If no alternative sources of money are found—or if those available are not sufficient to get projects to critical scale—then those projects may have to stop or progress much more slowly. In some cases, innovations never scale beyond the initial stages of development, regardless of their technical sophistication or desirability.

For a firm to get through this valley of death, it must either secure investors who believe in the innovation's potential or attract enough customers to sustain operations. True commercial viability typically requires reducing per-unit costs to an affordable level for most customers. This can be particularly challenging for projects that require very large capital investments.

Bridge funding, which could come from government entities, banks, or other sources, may help to establish commercial viability, but it is an ongoing challenge to distinguish between genuinely promising innovations and those that appear to be innovative but are not commercially viable. Firms failing to cross the valley of death could be acquired by foreign competitors from China and other nations with a greater willingness to invest in a technology not yet proven in the marketplace.

Focused research organizations (FROs) are a new nonprofit funding model designed to bridge the valley of death by providing financial support to teams of scientists and engineers for rapid prototyping and testing of technologies that advance the public good. Convergent Research, a nonprofit established in 2021 to support FROs, received $50 million in philanthropic donations in March 2023 to start two new FROs.[56]

Infrastructure for Innovation

KEY TAKEAWAYS

- Standards enable interoperability, lower costs, and support global trade, but they can also stifle innovation and be manipulated for market control or geopolitical advantage.
- Manufacturing is vital for economic resilience and security, especially amid global supply chain disruptions and strategic competition with China and other nations. Technological advances like robotics and AI are reshaping production, while policies such as the CHIPS and Science Act of 2022 aim to boost domestic capacity.
- Cybersecurity protects data, systems, and intellectual property from threats, ensuring research integrity and confidentiality. However, maintaining robust security can conflict with the open culture of research environments.

Standards

Standards are agreements—often formal ones—that specify technical or other requirements for products, processes, or services. Their primary function is to ensure that different systems and components are interoperable (i.e., they can work together effectively). Examples include standardized shipping containers, which revolutionized global logistics, and universal information technology protocols, like Universal Serial Bus (USB) and internet-related standards, which facilitate a high degree of compatibility across devices and networks.

Standards play a key role in enabling the diffusion of new technologies and are a foundational element of modern economies. They provide common frameworks that facilitate interoperability, compatibility, and safety, which are essential for scaling innovations from isolated prototypes to widespread adoption. However, while standards offer significant benefits, they also present challenges, including potential constraints on innovation, market power imbalances, and geopolitical complexities.

Research has shown that standards lower transaction costs, reduce uncertainty for producers and consumers, and enable the creation of large, interconnected markets.[57] Also, by codifying knowledge and best practices through consensus-based processes, they often play an important role in transforming scientific discoveries into commercial technologies, products, and services.[58] Standards streamline coordination by minimizing ambiguity in performance expectations and by supporting interoperability, which in turn accelerates market uptake.

Another key function of standards is to foster trust. Quality and safety standards mitigate the risks associated with innovative products, reducing uncertainty and bridging information gaps between developers and users.[59] Such trust facilitates early adoption and can contribute to the success of new technologies in the marketplace. For example, in the 1980s and 1990s, Europe's early adoption of the Global System for Mobile Communications (GSM) standard enabled rapid technical development, swift deployment, and seamless cross-border roaming, helping Europe quickly become a global telecom leader.

Overall, the economic impact of standards is substantial. Empirical studies across countries such as Germany, France, Australia, the United Kingdom, and Canada estimate that standards contribute between 0.2 and 0.9 percent to annual GDP growth.[60] They can also play a strategic role in national competitiveness. Countries that actively participate in international standard-setting bodies can influence the direction of technological development and ensure that their domestic industries are well positioned in global markets.

While they provide many benefits, standards also present challenges, including potential constraints on innovation, market power imbalances, and geopolitical complexities. One often-expressed concern is that prematurely deploying a set of standards may stifle innovation by locking in a particular technology or approach, making it difficult for newer, radically different, and potentially superior solutions to gain traction.

For example, the widespread use of the QWERTY keyboard—originally designed for mechanical typewriters in the nineteenth century—continues despite well-documented evidence that alternative layouts are significantly easier to learn and allow faster typing. The main reason the QWERTY layout remains dominant is that switching to another layout is seen as too costly for individuals and organizations.

Furthermore, the standardization process is often time and resource intensive, and dominant firms may use their influence to ensure standards favor their proprietary technologies, raising rivals' costs and creating barriers to entry. This can lead to market concentration, reduced competition, and the risk of industries becoming locked into aging solutions.

This history of "standards wars" between incompatible technologies illustrates how, when no clear standard prevails, competing standards can fragment markets and slow global diffusion.[61] The videotape format war between VHS (Video Home System) and Betamax in the late 1970s into the 1980s is a classic example. Betamax had better picture quality but shorter recording times, higher costs, and restrictive licensing. VHS offered longer recording times, lower prices, and open licensing, attracting more manufacturers and broader studio support. For years, both formats coexisted, forcing consumers and retailers into incompatible ecosystems. VHS's advantages eventually secured dominance, thus forcing Betamax users to convert to VHS and to lose their original investments.[62]

Finally, standardization is increasingly entangled with global geopolitical competition, as countries vie for influence in international standards bodies to shape rules that favor domestic firms. This can lead to the emergence of competing standards regimes, undermining global interoperability and raising the complexity and cost of doing business internationally.

Manufacturing

Manufacturing plays an increasingly critical role in the US economy and national security, driven by a rapidly evolving geopolitical landscape and the growing recognition of vulnerabilities in global supply chains. Strategic competition with China has intensified US

This history of "standards wars" between incompatible technologies illustrates how . . . competing standards can fragment markets and slow global diffusion.

concerns over economic security and technological leadership, especially as China advances in semiconductors, AI, and clean energy. Partly in response, the United States is prioritizing domestic manufacturing to reduce reliance on foreign sources and to better protect critical technologies.[63]

Recent global events, such as the COVID-19 pandemic and increased use of export restrictions, have exposed the fragility of international supply chains and revealed how disruptions abroad can have immediate and severe impacts on the availability of essential goods in the United States. This has led policymakers and industry leaders to focus on expanding domestic production capacity and achieving greater technological self-sufficiency.

On the supply side, technology innovation is driving a manufacturing renaissance. Advances in robotics, AI, additive manufacturing (3-D printing), advanced materials (see chapter 5, on materials science), and big data analytics are transforming how goods are designed, produced, and delivered. These technologies enable greater customization, faster prototyping, and more efficient production. Automation and AI may also offset labor cost advantages that previously favored offshoring.

Manufacturing is closely linked to national security. A strong domestic manufacturing base would ensure that the United States can produce critical defense systems at significant scale, maintain technological superiority, and respond to emerging threats. It would also reduce risks such as espionage, intellectual property theft, and supply chain subversion that are often associated with foreign manufacturers. Additionally, manufacturing supports millions of jobs, drives innovation, and stabilizes supply chains across the economy.

Revitalizing US manufacturing is a prospect that enjoys bipartisan support. For example, under the Biden administration, the CHIPS and Science Act was passed in 2022. It was intended to return a significant amount of chip fabrication to American shores. Similarly, the Trump administration's Made in America Manufacturing Initiative is intended to encourage and enable domestic manufacturers—especially small and midsize firms—to become preferred suppliers for government contracts, particularly in critical sectors like aerospace, defense, and energy.

Cybersecurity

Cybersecurity encompasses the technologies, processes, and policies that protect computer systems, networks, and the information they contain from malicious activities by adversaries or unscrupulous actors. The field centers on the protection of three core principles: confidentiality, integrity, and availability. Confidentiality ensures data privacy, preventing unauthorized disclosure. Integrity maintains data and program accuracy, guarding against unauthorized alterations. Availability ensures data and computing resources are accessible to authorized users, especially during critical times.

Initially, cybersecurity as a technical discipline focused on secure programming languages and robust software architectures, which created systems more resistant to threats like malware and advanced cyberattacks. As the internet expanded, and networked devices proliferated, the scope of cybersecurity broadened to include the protection of infrastructure that supports data transmission and storage. The field has since evolved to address new challenges, including social engineering attacks, digital misinformation, information warfare, and the risks posed by AI at both the human and system levels. (The development and use of foundational AI models introduce additional cybersecurity risks, as discussed in chapter 3, on cryptography and computer security.)

As a national-level issue, cybersecurity policy measures are often associated with private-sector businesses and government. But cybersecurity is also a critical concern for R&D in academia and industry. In research settings, one major concern is ensuring the integrity of scientific data. The deletion, destruction, or subtle alteration of research data can undermine scientific progress by wasting resources or undermining the validity of results. Computer programs used in research are similarly vulnerable; minor, undetected changes in them can call into question the accuracy of previously collected or analyzed data.

Protecting the confidentiality of work products—such as datasets and draft working papers—is equally important. Unauthorized access to confidential datasets can breach agreements and compromise academic integrity, while premature disclosure of draft research can undermine claims of priority and reveal incomplete or inconsistent findings. Computers that manage laboratory data collection are susceptible to attacks that could disrupt research continuity by corrupting data or damaging equipment.

While technical safeguards exist to address these cybersecurity challenges, maintaining them in academic settings requires substantial management effort. The informal, collegial, and flexible culture common in research labs often views rigorous security practices as disruptive, which can lead to resistance to them or to inconsistent implementation.

Another growing threat involves the selective targeting of personnel working on key research projects. Researchers may face cyber harassment, financial compromise, or threats to family members. Attacks on professional ethics through social media or online forums can damage reputations, cause personal distress, and reduce productivity.

Addressing these multifaceted cybersecurity challenges requires a careful balance between robust protection and the need for open, collaborative research environments. Effective cybersecurity policies and practices must be adaptable, recognizing the unique risks and cultural dynamics present in academic and research settings while ensuring the integrity, confidentiality, and availability of critical information and systems.

NOTES

1. See, for example, Josh Luckenbaugh, "Just In: U.S. Falling Behind China in Critical Tech Race, Report Finds," *National Defense*, July 17, 2023, https://www.nationaldefensemagazine.org/articles/2023/7/17/us-falling-behind-china-in-critical-tech-race-report-finds; Jeremy Neufeld, *STEM Immigration Is Critical to American National Security* (Institute for Progress, 2022), https://ifp.org/stem-immigration-is-critical-to-american-national-security/; Craig Cohen and Alexander Kisling, eds., "Part II: Winning the Economic and Tech Race," in *2024 Global Forecast: A World Dividing* (Center for Strategic and International Studies, 2024), https://csis-website-prod.s3.amazonaws.com/s3fs-public/2024-01/240130_GlobalForecast_2024_WinningEconomic_TechRace.pdf.

2. International Monetary Fund, "Is Productivity Growth Shared in a Globalized Economy?," chap. 4 in *World Economic Outlook 2018: Cyclical Upswing, Structural Change* (April 2018), https://www.imf.org/en/Publications/WEO/Issues/2018/03/20/world-economic-outlook-april-2018.

3. Catherine Tucker, "Network Effects and Market Power: What Have We Learned in the Last Decade?," *Antitrust*, 2018, 72–9.

4. See, for example, Manuel Hoffmann, Frank Nagle, and Yanuo Zhou, "The Value of Open Source Software," Harvard Business School Strategy Unit Working Paper no. 24-038, https://doi.org/10.2139/ssrn.4693148.

5. Davide Rigo, "Global Value Chains and Technology Transfer: New Evidence From Developing Countries," *Review of World Economics* 157 (2021): 271–94, https://doi.org/10.1007/s10290-020-00398-8.

6. Commission on the Theft of American Intellectual Property, "IP Commission 2021 Review: Updated Recommendations" (2021), https://www.nbr.org/wp-content/uploads/pdfs/publications/ip_commission_2021_recommendations_mar2021.pdf.

7. Pamela Samuelson and Suzanne Scotchmer, "The Law and Economics of Reverse Engineering," *Yale Law Journal* 111, no. 7 (2002):1575–1663, https://www.yalelawjournal.org/article/the-law-and-economics-of-reverse-engineering.

8. Adam Satariano et al., "Elon Musk's Unmatched Power in the Stars," *New York Times*, July 28, 2023, https://www.nytimes.com/interactive/2023/07/28/business/starlink.html.

9. On solar panels, see François Lafond, Aimee Gotway Bailey, Jan David Bakker, et al., "How Well Do Experience Curves Predict Technological Progress? A Method for Making Distributional Forecasts," *Technological Forecasting and Social Change* 128 (2018): 104–17, https://doi.org/10.1016/j.techfore.2017.11.001. On LEDs, see Brian F. Gerke, Allison T. Ngo, Andrea L. Alstone, and Kibret S. Fisseha, *The Evolving Price of Household LED Lamps: Recent Trends and Historical Comparisons for the US Market*, Office of Science and Technical Information (US Department of Energy, 2014), https://doi.org/10.2172/1163956.

10. "A Walk Through Time—A Revolution in Timekeeping," National Institute of Standards and Technology, August 12, 2009, https://www.nist.gov/pml/time-and-frequency-division/popular-links/walk-through-time/walk-through-time-revolution.

11. Keith Fuglie, Madhur Gautam, Aparajita Goyal, and William F. Maloney, *Harvesting Prosperity: Technology and Productivity Growth in Agriculture* (World Bank, 2020), https://openknowledge.worldbank.org/bitstreams/3621191c-15f3-5ede-a89c-f7190d7e1dba/download.

12. "History of the Web," World Wide Web Foundation, October 18, 2009, https://webfoundation.org/about/vision/history-of-the-web/.

13. Amirhosein Toosi, Andrea Bottino, Babak Saboury, Eliot Siegel, and Arman Rahmim, "A Brief History of AI: How to Prevent Another Winter (a Critical Review)," *PET Clinics* 16, no. 4 (2021): 449–69, https://doi.org/10.1016/j.cpet.2021.07.001.

14. See, for example, Vaclav Smil, "Techno-Optimism, Exaggerations and Realistic Expectations," in *Invention and Innovation: A Brief History of Hype and Failure* (MIT Press, 2023).

15. University of Utah, "'Simple Experiment' Results in Sustained N-Fusion at Room Temperature for First Time," news release, March 23, 1989, archived at https://newenergytimes.com/v2/archives/UofUtah/19890323-Univ-Utah-Press-Release.pdf.

16. For example, Ira Magaziner, paraphrased by author Gary Taubes in Gary Taubes, *Bad Science: The Short Life and Weird Times of Cold Fusion* (Random House, 1993).

17. Stephan Wilkinson, "How the AK-47 Became the 'Weapon of the Century," *Military Times*, December 12, 2017, https://www.militarytimes.com/off-duty/gearscout/2017/12/12/how-the-ak-47-became-the-weapon-of-the-century/.

18. Audrey Cronin, *Power to the People: How Open Technological Innovation Is Arming Tomorrow's Terrorists* (Oxford University Press, 2020).

19. Josh Abramson, Jonas Adler, Jack Dunger, et al., "Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3," *Nature* 630 (2024): 493–500, https://doi.org/10.1038/s41586-024-07487-w.

20. This report, chapter 5, on materials science.

21. Po-Wen Chan and Balaji Chandrasekaran, "Materials Engineering: The True Hero of Energy-Efficient Chip Performance," *Applied Materials*, August 1, 2024, https://www.appliedmaterials.com/us/en/blog/blog-posts/the-true-hero-of-energy-efficient-chip-performance.html.

22. Frontiers in Space Technology, "Next Generation of Materials for Space Applications," *Frontiers*, research topic, accessed September 18, 2024, https://www.frontiersin.org/research-topics/55679/next-generation-of-materials-for-space-applications.

23. Yongli Qi, Seung-Kyun Kang, and Hui Fang, "Advanced Materials for Implantable Neuroelectronics," *MRS Bulletin* 48 (2023): 475–83, https://doi.org/10.1557/s43577-023-00540-5.

24. Anil D. Pathak, Shalakha Saha, Vikram Kishore Bharti, et al., "A Review on Battery Technology for Space Application," *Journal of Energy Storage* 61 (2023): 106792, https://doi.org/10.1016/j.est.2023.106792.

25. Orlando Burgos-Morales, M. Gueye, Laurie Lacombe, et al., "Synthetic Biology as Driver for the Biologization of Materials Sciences," *Materials Today Bio* 11 (2021): 100115, https://doi.org/10.1016/j.mtbio.2021.100115.

26. James M. Heather and Benjamin Chain, "The Sequence of Sequencers: The History of Sequencing DNA," *Genomics* 107, no. 1 (2016): 1–8, https://doi.org/10.1016/j.ygeno.2015.11.003.

27. A report from the National Research Council of the US National Academies of Sciences, Engineering, and Medicine addresses this distinction in more detail, drawing distinctions between "foundational technologies" and "technology applications," with the latter focusing on solving specific problems through an artful blend of technologies. Space and robotics are examples of technology applications that build on foundational advances in many fields. See Jean-Lou Chameau, William F. Ballhaus, and Herbert Lin, eds., *Emerging and Readily Available Technologies and National Security: A Framework for Addressing Ethical, Legal, and Societal Issues* (National Academies Press, 2014), https://doi.org/10.17226/18512.

28. See, for example, Maja Bacovic, Zivko Andrijasevic, and Bojan Pejovic, "STEM Education and Growth in Europe," *Journal of the Knowledge Economy* 13 (2022): 2348–71, https://doi.org/10.1007/s13132-021-00817-7.

29. US Department of Defense, *Industrial Capabilities Report to Congress: 2020 Annual Report* (January 2021), https://www.businessdefense.gov/docs/resources/USA002573-20_ICR_2020_Web.pdf.

30. Eric Chewning, Matt Schrimper, Andy Voelker, and Brooke Weddle, "Debugging the Software Talent Gap in Aerospace and Defense," McKinsey & Company, July 11, 2022, https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/debugging-the-software-talent-gap-in-aerospace-and-defense.

31. Wilson Miles and Jeremy Neufeld, "U.S. Needs International Talent to Maintain Tech Leadership," *National Defense Magazine*, April 14, 2025, https://www.nationaldefensemagazine.org/articles/2025/4/14/us-needs-international-talent-to-maintain-tech-leadership.

32. Gabrielle Athanasia and Jillian Cota, "The U.S. Should Strengthen STEM Education to Remain Globally Competitive," Center for Strategic and International Studies, April 1, 2022, https://www.csis.org/blogs/perspectives-innovation/us-should-strengthen-stem-education-remain-globally-competitive; Darrell M. West, "Improving Workforce Development and STEM Education to Preserve America's Innovation Edge," Brookings Institution, July 26, 2023, https://www.brookings.edu/articles/improving-workforce-development-and-stem-education-to-preserve-americas-innovation-edge/; Julia Yoon, "Innovation Lightbulb: Strengthening K–12 STEM Education for a Robust U.S. Technology Workforce," Center for Strategic and International Studies, June 28, 2024, https://www.csis.org/analysis/innovation-lightbulb-strengthening-k-12-stem-education-robust-us-technology-workforce.

33. National Science Board, *Talent Is the Treasure: Who Are We Leaving on the Bench?*, report no. NSB-2024-11 (March 2024), https://www.nsf.gov/nsb/publications/2024/2024_policy_brief.pdf.

34. National Science Board, *Talent Is the Treasure*; Yoon, "Innovation Lightbulb."

35. Grace Zichelli, "New Issue Brief: 2024 NAEP Results Are a Wake-up Call for American Education," Manhattan Institute, news release, May 15, 2025, https://manhattan.institute/article/new-issue-brief-2024-naep-results-are-a-wake-up-call-for-american-education.

36. Athanasia and Cota, "The U.S. Should Strengthen STEM Education"; Steven Deitz and Christina Freyman, *The State of U.S. Science and Engineering 2024*, National Science Board, Science and Engineering Indicators (National Science Foundation, 2024), https://ncses.nsf.gov/pubs/nsb20243/; Yoon, "Innovation Lightbulb"; National Science Board, *Talent Is the Treasure*.

37. Table 1.1 in *The ACT: Profile Report—National, Graduating Class 2024*, ACT, https://www.act.org/content/dam/act/unsecured/documents/2024-act-national-graduating-class-profile-report.pdf.

38. Program for International Student Assessment, "PISA 2022 Mathematics Literacy Results," Institute of Education Sciences,

National Center for Education Statistics, accessed September 18, 2024, https://nces.ed.gov/surveys/pisa/pisa2022/mathematics/international-comparisons/.

39. Tuan D. Nguyen, "The Supply and Quality of STEM Teachers," *Humanities and Social Sciences Communications* 12, no. 1 (2025), https://doi.org/10.1057/s41599-025-04648-8. Though published in 2025, this paper cites other work whose original data provenance dates to 2011–12.

40. Michael Marder, "How Bad Is the U.S. STEM Teacher Shortage?" *Medium*, December 8, 2021, https://uteachinstitute.medium.com/u-s-stem-teacher-shortages-in-2017-2018-1d8314a93ba1.

41. Marder, "How Bad Is the U.S. STEM Teacher Shortage?"

42. Michael Polyani, *Personal Knowledge* (University of Chicago Press, 1962); Thomas S. Kuhn, *Structure of Scientific Revolutions*, 2nd ed. (University of Chicago Press, 1970); for more recent discussions, see: Vincent-Wayne Mitchell, William S. Harvey, and Geoffrey Wood, "Where Does All the 'Know How' Go? The Role of Tacit Knowledge in Research Impact," *Higher Education Research & Development* 41, no. 5 (2022): 1664–78, https://doi.org/10.1080/07294360.2021.1937066; Tim Thornton, "Tacit Knowledge as the Unifying Factor in Evidence Based Medicine and Clinical Judgement," *Philosophy, Ethics, and Humanities in Medicine* 1, no. 2 (2006), https://doi.org/10.1186/1747-5341-1-2.

43. Shai Bernstein, Rebecca Diamond, Abhisit Jiranaphawiboon, Beatriz Pousada, and Timothy McQuade, "The Contribution of High-Skilled Immigrants to Innovation in the United States," Research Briefs in Economic Policy no. 350, Cato Institute, September 20, 2023, https://www.cato.org/research-briefs-economic-policy/contribution-high-skilled-immigrants-innovation-united-states; Jennifer Hunt and Marjolaine Gauthier-Loiselle, "How Much Does Immigration Boost Innovation?," *American Economic Journal: Macroeconomics* 2, no. 2 (2010): 31–56, https://doi.org/10.1257/mac.2.2.31.

44. Dany Bahar et al., "*Talent Flows and the Geography of Knowledge Production: Causal Evidence from Multinational Firms*," Harvard Business School Technology and Operations Mgmt. Unit Working Paper no. 22-047 (December 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4005693.

45. Immigration, Refugees, and Citizenship Canada, "Canada's Tech Talent Strategy," Government of Canada, last modified June 27, 2023, https://www.canada.ca/en/immigration-refugees-citizenship/news/2023/06/canadas-tech-talent-strategy.html.

46. Britta Glennon and David R. Dollar, *Dollar & Sense*, podcast, "What's Behind the Globalization of R&D?," Brookings Institution, April 26, 2021, MP3 Audio, 25:02, https://www.brookings.edu/articles/whats-behind-the-globalization-of-rd.

47. National Academies of Sciences, Engineering, and Medicine, *Information Technology Innovation: Resurgence, Confluence, and Continuing Impact* (National Academies Press, 2020), https://doi.org/10.17226/25961.

48. Neil C. Thompson, Shuning Ge, and Yash M. Sherry, "Building the Algorithm Commons: Who Discovered the Algorithms That Underpin Computing in the Modern Enterprise?," *Global Strategy Journal* 11, no. 1 (2021): 17–33, https://doi.org/10.1002/gsj.1393.

49. National Research Council, *Research Universities and the Future of America: Ten Breakthrough Actions Vital to Our Nation's Prosperity and Security* (National Academies Press, 2012), https://doi.org/10.17226/13396.

50. Graham Andrews, "New Data Show Universities Are Increasing R&D Activity," Association of American Universities, December 6, 2024, https://www.aau.edu/newsroom/leading-research-universities-reportnew-data-show-universities-are-increasing-rd-activity.

51. James Manyika and William McRaven, *Innovation and National Security: Keeping Our Edge*, Independent Task Force Report, no. 77 (Council on Foreign Relations, 2019), 21, https://www.cfr.org/report/keeping-our-edge/pdf/TFR_Innovation_Strategy.pdf.

52. Andrews, "New Data Show Universities Are Increasing R&D Activity."

53. National Center for Science and Engineering Statistics, "Long-Term Trends Show Decline in Federally Funded R&D as a Share of GDP While Business-Funded R&D Increases as a Share of GDP," NSF 25-334 (US National Science Foundation, 2025), https://ncses.nsf.gov/pubs/nsf25334.

54. James Pethokoukis, "Broken Promises: CHIPS Act Funding for Science Research Falls Short," *Faster, Please!*, American Enterprise Institute, March 13, 2024, https://www.aei.org/articles/broken-promises-chips-act-funding-for-science-research-falls-short/.

55. Organisation for Economic Co-operation and Development, "OECD Data Explorer: Gross Domestic Expenditure on R&D by Sector of Performance and Type of R&D," 2024, https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CScience%252C%20technology%20and%20innovation%23INT%23%7CResearch%20and%20development%20%28R%26D%29%23INT_RD%23&pg=0&fc=Topic&bp=true&snb=18&vw=tb&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_RDS_GERD%40DF_GERD_TORD&df[ag]=OECD.STI.STP&df[vs]=1.0&dq=CHN%2BUSA.A.._T...BR...USD_PPP.Q&pd=2012%2C2021&to[TIME_PERIOD]=false. Analysis conducted by comparing two methods of measuring annual increases of US and China spending on basic research and projecting rates into future.

56. Alex Knapp, "Why Billionaires Ken Griffing and Eric Schmidt Are Spending $50 Million on a New Kind of Scientific Research," *Forbes*, March 17, 2023, https://www.forbes.com/sites/alexknapp/2023/03/17/why-billionaires-ken-griffin-and-eric-schmidt-are-spending-50-million-on-a-new-kind-of-scientific-research.

57. G. M. Peter Swann, *The Economics of Standardization: An Update*, report for the UK Department of Business, Innovation and Skills (Innovative Economics Limited, 2010), https://assets.publishing.service.gov.uk/media/5a790abd40f0b679c0a0812d/10-1135-economics-of-standardization-update.pdf.

58. European Commission, "A strategic vision for European standards: Moving forward to enhance and accelerate the sustainable growth of the European economy by 2020" (Communication) COM (2011) 0311 final, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0311:FIN:en:PDF.

59. *Building Trust: The Conformity Assessment Toolbox*, International Organization for Standards, 2023, https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/casco_building-trust.pdf.

60. Andre Jungmittag and Axel Mangelsdorf, *The Economic Benefits of Standardization*, DIN German Institute for Standardization, 2011, https://www.din.de/resource/blob/89552/68849fab0eeeaafb56c5a3ffee9959c5/economic-benefits-of-standardization-en-data.pdf.

61. Carl Shapiro and Hal R. Varian, "The Art of Standards Wars," *California Management Review* 41, no. 2 (1999): 8–32, https://doi.org/10.2307/41165984.

62. Brad Kelechava, "VHS vs Betamax: Standard Format War," *The ANSI Blog*, May 5, 2016, https://blog.ansi.org/ansi/vhs-vs-betamax-standard-format-war/.

63. "The CHIPS Act: What It Means for the Semiconductor Ecosystem," PricewaterhouseCoopers, n.d., https://www.pwc.com/us/en/library/chips-act.html.

STANFORD EXPERT CONTRIBUTOR

Dr. Herbert S. Lin
SETR Director and Editor in Chief, Research Fellow at the Hoover Institution

12

TECHNOLOGY APPLICATIONS BY POLICY AREA

This chapter explores applications from each technology field described in the report as they relate to five important policy themes: economic growth, national security, environmental and energy sustainability, health and medicine, and civil society.

Economic Growth

Artificial intelligence (AI) AI may significantly boost productivity across many sectors of the economy. Large language models such as ChatGPT have already demonstrated how they can be used in a variety of diverse fields, including law, customer support, computer programming, and journalism. Generative AI, a form of AI that creates new text, images, and other content, is expected to raise global GDP by $7 trillion and lift productivity growth by 1.5 percent over a ten-year period, if adopted widely.

Biotechnology and synthetic biology Biotechnology is poised to emerge as a general-purpose technology that can be applied broadly, with the capacity to revolutionize areas such as healthcare and manufacturing. Biological processes could ultimately produce as much as 60 percent of the physical inputs to the global economy. Already, biotechnology and synthetic biology are enablers for advances in medicine and healthcare (e.g., vaccines and cancer treatments), agriculture (e.g., drought-resistant crops), food (e.g., nutritionally enriched vegetables), and energy production (e.g., biofuels). Potential applications also include biotic semiconductors, magnets, fiber optics, and data storage.

Cryptography and computer security Blockchain technologies can effectively provide provenance in supply chains as well as personal identity management that curbs fraud and identity theft, leading to more secure and efficient transactions. Blockchain technology also underpins cryptocurrencies. A US central bank digital currency, or CBDC, a form of digital currency that does not necessarily use blockchains, could help reduce inefficiencies in US deposit markets, promoting broader participation in the financial system.

Energy technologies The Accelerating Deployment of Versatile, Advanced Nuclear for Clean Energy (ADVANCE) Act of 2024 sought to strengthen US global leadership on nuclear energy by directing the Nuclear Regulatory Commission (NRC) to coordinate international nuclear export licensing and to establish an International Nuclear Reactor Export and Innovation Branch of the NRC's International Programs office. With growing demand around the world for nuclear power, there are many opportunities expected for the US reactor industry to export its products.

Materials science Lighter and stronger materials will increase the energy efficiency of vehicles used to transport people and cargo. New semiconductor materials enable new types of chips and other information processing hardware. Technological innovations are also offering new ways to produce low-carbon steel and cement.

Neuroscience Interventions for those with neural disorders include pharmaceuticals that curb, treat, or reverse neurodegenerative conditions; diagnostics to identify early onset of such conditions; and rehabilitation therapies that help people suffering from them engage in the activities of daily living. By helping to address neurogenerative diseases more effectively, research in the field could allow people to remain in the workforce longer and be more productive, as well as reduce the burden on caregivers, who often need to take time off work to look after relatives and friends.

Quantum technologies Quantum computing can address problems in portfolio optimization, modeling for drug discovery, and the improvement of delivery routes. Quantum sensing may be important for subsurface exploration for oil and minerals and quality control in semiconductor manufacture.

Robotics Robots are used widely today, including in manufacturing; on-demand delivery services; surgery; science and exploration; food production; disaster assistance; security and military services; and transportation. Innovations in robotics have enormous potential to increase productivity in many fields and perhaps to create new types of jobs. But robots that involve physical labor and presence may also eliminate some jobs and change others, creating the need for retraining people and other measures to address short-term impacts.

Semiconductors Semiconductors are an enabling technology for any application that can be improved through the use of information. They provide the computing capabilities that many sectors of the economy rely on. As such, they are key drivers of economic activity and growth. However, reductions in the cost of semiconductors and increases in processing power are likely to become less frequent or regular in the future—and predictions about future economic growth attributable to improvements in semiconductor technology may prove to be overly optimistic.

Space Space activities play critical roles in our daily lives and the economy, from enabling global navigation systems to providing precise time information for financial transactions. Expanding commercial activities are expected to drive high growth in the space sector. In the future, through things such as asteroid mining and space-based power production, space activities could become even bigger drivers of economic growth on Earth.

National Security

Artificial intelligence Because AI enables more rapid processing of an expanded range of data inputs, all aspects of military operations potentially benefit from it. Possible applications include managing military logistics; improving the effectiveness and efficiency of maintaining equipment; managing electronic medical records; navigating autonomous vehicles; operating drone swarms; recognizing targets; performing intelligence analysis; developing options for command decisions; and enhancing war gaming to develop and refine plans. However, the US Department of Defense's ethical considerations for the development and deployment of AI capabilities (especially in nuclear command and control) may not be shared by adversaries.

Biotechnology and synthetic biology With synthetic biology becoming increasingly available to state and nonstate actors, there are concerns that a malicious actor could create or deploy weaponized organisms or threaten the provision of biologically developed foods, medicines, fuels, and other products to coerce others. Conversely, the prospect of distributed biomanufacturing offers possibilities for localized biodefense and a larger degree of independence from foreign suppliers of many raw materials. China is investing considerably more resources in biotechnology than the United States, creating the potential for a Sputnik-like strategic surprise.

Cryptography and computer security Adversaries are likely to have been storing encrypted data, hoping that future advances in quantum computing and other digital capabilities will allow them to crack the encryption protecting the information. Efforts are already underway to create new encryption methods that would be quantum resistant. Separately, zero-knowledge proof methodology to cooperatively track and verify numbers of tactical nuclear warheads may benefit future arms control agreements.

Energy technologies The United States is no longer the world leader in energy manufacturing at scale. For instance, China and other countries with lower operating costs control most of the manufacturing, supply chain, and critical minerals for battery and solar cell production. US energy security will require expansion of domestic production and manufacturing, as well as collaboration with allies and partners to better protect energy supply chains. Moreover, there are concerns that a global increase in fission reactors will result in a greater risk of nuclear proliferation (i.e., the spread of nuclear weapons), especially to nonnuclear states or nonstate actors. However, some believe that the emissions-free potential of fission reactors is worth the risk of proliferation, which can be minimized through carefully implemented safeguards. Fuel security for nuclear power remains an issue as well—America currently imports more than 90 percent of its uranium, with about half coming from Kazakhstan and Russia.

Materials science Improvements in materials science and nanotechnology can advance capabilities in stealth technology, camouflage, and body armor and can increase the energy content in explosives. Quantum dots—materials that are smaller than about 100 nanometers in all dimensions—can be used in sensors for detecting agents associated with chemical and biological warfare.

Neuroscience Neuroscience may help illuminate the nature of traumatic brain injuries and post-traumatic stress disorder, thereby leading to better treatments for these conditions. Brain–machine interfaces could also enable new prostheses for wounded combatants.

Quantum technologies Quantum inertial sensing can provide precise timing and position information in GPS-denied environments (i.e., places where global positioning systems are not available). Quantum magnetometers can enable detection and tracking of submarines, camouflaged weapons, and mines by sensing small magnetic anomalies from long distances. Quantum imaging technologies like quantum LIDAR (light detection and ranging) may enable better vision through obscurants such as smoke, fog, or foliage and enhance detection of hidden targets with high resolution.

Robotics Advances in robotics can assist military forces with the transportation of equipment and supplies, urban warfare, autonomous vehicle deployment, and search-and-rescue efforts. Additionally, robotics can assist with mine clearance, disaster recovery, and firefighting. Some military robots, such as lethal autonomous weapons systems, raise questions of robo-ethics on the battlefield. Given the pressure for militaries to act more rapidly, many observers believe that decisions of lethal force will be turned over to computers, while others insist that life-and-death decisions must remain with humans.

Semiconductors Modern military hardware is critically dependent on semiconductor technology for information processing. The primary fabricator of semiconductor chips globally is Taiwan. Taiwan is home to two of the three leading manufacturers: the Taiwan Semiconductor Manufacturing Company and the United Microelectronics Corporation. China's long-held interest in reunification with Taiwan and its rising military capabilities and assertiveness toward Taiwan are raising deep concerns in regard to semiconductors. Many are concerned about the potential for a Chinese blockade or other actions that could disrupt the global semiconductor supply chain and raise the risk of military conflict between the United States and China. The Creating Helpful Incentives to Produce Semiconductors (CHIPS) and Science Act of 2022 is intended to reduce the risk of supply chain disruption, but major initiatives called for in the legislation have not been fully funded.

Space Communications, surveillance, and navigation in denied areas are essential functions for military forces. In the future, nonnuclear weapons may be based in space and used to attack terrestrial and space targets. Satellites are also essential for the detection of launched ballistic missiles, nuclear weapons explosions, and electromagnetic emissions from other nations. The emergence of low-cost, high-quality information from space-based assets that are largely commercial has been a driver of open-source intelligence (OSINT). Unclassified intelligence like OSINT has the potential to upend traditional intelligence processes built on classified information collection and analysis. The net effect of OSINT could be a declining US intelligence advantage, as more countries, organizations, and individuals can collect, analyze, and disseminate high-quality intelligence without expensive, space-based government satellite capabilities. The commercialization of space also puts powerful capabilities in the hands of individuals and organizations who are not accountable to voters and whose interests may not be aligned with those of the US government.

Environmental and Energy Sustainability

Artificial intelligence AI capabilities can greatly improve global sustainability efforts, from helping farmers identify which produce or livestock are appropriate to harvest, to helping analyze weather patterns to prepare populations and infrastructure for extreme or unusual conditions. At the same time, training and using AI models requires a large amount of energy, and the energy demand to support these activities is expected to grow significantly in the future.

Biotechnology and synthetic biology Synthetic biology can contribute to new methods for energy production and environmental cleanup. Electrobiosynthesis is a biotechnology that enables plant-free bioproduction in places where soils are poor, water is scarce, or climate and weather are too variable to support traditional agriculture.

Cryptography and computer security Blockchain technologies can provide a transparent and secure way to track the movement of goods. This includes tracking their origin, quantity, and other relevant information, thereby improving efficiency in global supply chains and limiting illegal extractions of certain materials. Although some established cryptocurrencies, such as Bitcoin, require massive amounts of energy, newer cryptocurrencies require far less.

Energy technologies New investments in energy research and development are enabling advances in clean electricity generation, long-distance transmission lines, lighting based on light-emitting diodes (LEDs), and electric car batteries. Long-duration energy storage is a critical field for climate and sustainability goals. The development of batteries for electric grids that can store energy for weeks or months is needed to support the use of solar and other intermittent renewable energy sources. Renewable fuels, especially hydrogen, can replace hydrocarbons in transportation and industry. However, new hydrogen production and storage methods are needed to make its use cost-effective at scale. Nuclear power could help the United States reach sustainability goals. However, it is unclear whether enough reactors can become operational in time to meet commitments to triple nuclear generation of electricity by 2050 compared to the 2020 baseline. Moreover, nuclear waste remains an environmental policy issue, and the United States has no enduring plan for a long-term storage solution.

Materials science Innovations in materials science and engineering are creating new and sustainable plastics that are easier to recycle. New materials can also advance the electrification of transportation and industry, which is integral to decarbonization strategies. They can also support the design of relatively cheap batteries that last a long time and can be quickly recharged. Nanomaterials such as quantum dots can further improve the efficiency of solar cells and biodegradable plastics. However, some innovations in the field have potential downsides, too. For instance, the long-term dangers of nanoparticles released into the environment at the end of their life cycle are unknown.

Neuroscience Sustainability on a planet with finite resources requires that decision makers and the people they represent are able to make trade-offs between immediate rewards and future gains. Neuroscientists have found evidence for cognitive predisposition favoring short-term gains over long-term rewards, based on functional magnetic resonance imaging (fMRI) brain scans of people making choices between immediate and delayed reward.

Quantum technologies Because of their improved sensitivity, quantum sensors can provide more precise real-time monitoring of air, water, and soil pollution. Important use cases include detecting trace pollutants, greenhouse gases, and microplastics with high specificity.

Robotics The deployment of robots primarily for the "three D's"—jobs considered dull, dirty, or dangerous—enables robotic cleanup of environmentally hazardous materials and their operation in environments that can be dangerous for humans, such as nuclear reactors. Robots are also valuable in the construction, maintenance, and management of solar and wind farms.

Semiconductors Transitioning to renewable energy sources will require vast amounts of semiconductors. Advanced chips are integral to electric vehicles, solar arrays, and wind turbines. Design innovations will continue to improve the energy efficiency of chips.

Space Remote sensing data can create a "digital twin" of Earth to track and model environmental change and the movement of humans and animals, informing disaster response and sustainable development policies. The development of space technologies will help to address food security, greenhouse gas emissions, renewable energy, and supply chain optimization. Satellite imagery, combined with weather data and powered by predictive optimization algorithms, could increase crop yields. It could also detect greenhouse gas emissions to identify natural-gas leaks and verify compliance with regulations. Advancing space technologies could enable mining from the Moon and asteroids of minerals that are hard to find on Earth. It could also enable the transmission of sustainable solar energy directly to Earth from space.

Health and Medicine

Artificial intelligence AI data analytics are already improving the accuracy of healthcare assessments and procedures. Continued advancement could place AI-monitored cameras and sensors in the homes of elderly or at-risk patients to provide prompt attention in case of emergency while protecting patient privacy. AI-operated mobile robots can potentially replace basic nursing care.

Biotechnology and synthetic biology Synthetic biology has remarkable potential to contribute to the creation of new drugs as well as to pathogen detection and neutralization. It can also help to reduce disease transmission, personalize medicine through genetic modifications, improve cancer treatment, and offer custom lab-grown human tissue for medical testing. DNA sequencers and synthesizers using the internet allow researchers around the world to obtain information on viruses—and potentially vaccines or cures—faster than a pandemic can spread. However, that same speed and accessibility raise concerns about potential misuse of the technology by bad actors. It is also unclear how some new biological organisms will interact with the natural and human environments.

Cryptography and computer security Blockchain technology can securely store all data from a person's important documents, including medical records, in encrypted form while facilitating selective data retrieval that protects a patient's privacy. This approach enables the performance of data analytics on aggregated and anonymized datasets, enabling researchers and internal auditors to access information without violating patients' privacy rights.

Energy technologies A transition from fossil fuel energy to a renewable energy–based world economy would reduce greenhouse gas emissions and prevent thousands of premature deaths from pollution and extreme weather events. Eliminating energy-related air pollution in the United States alone could prevent more than fifty thousand deaths annually and save hundreds of billions of dollars a year from avoided illness. Reducing carbon dioxide emissions will result in less extreme climates, which in turn will lead to fewer health problems from extreme heat.

Materials science Materials science and nanotechnology are improving the capabilities and effectiveness of medical devices and the delivery of treatments. For example, wearable electronic devices made from flexible materials can conform to skin or tissues to provide specific sensing or actuating functions; devices like "electronic skin," or e-skin, can sense external stimuli such as temperature or pressure; and "smart bandages" with integrated sensors can significantly accelerate the healing of chronic wounds. Injectable hydrogels can fine-tune long-term delivery of medications, which can lead to improvements in the administration and efficacy of essential medicines such as insulin. Nanomaterials like quantum dots are being used as fluorescent markers in biological systems to improve the contrast of biomedical images. Finally, biosensors allow the rapid testing of blood for bacterial pathogens.

Neuroscience Advances in neuroscience may help address neurodegeneration and related diseases, such as chronic pain, depression, opioid dependency, and Alzheimer's disease. These advances could dramatically improve the quality of life of patients (and their families) and potentially reverse the anticipated rising costs associated with care. However, too many fundamental gaps still remain in our understanding of the brain for confidence in the rapid progress of treating such illnesses.

Quantum technologies Quantum sensors can provide sensitive detection of the small magnetic fields generated by neural activity. This enables the development of noninvasive, wearable systems that can perform three-dimensional mapping of brain activity in real time, with high spatial and temporal resolution, while subjects move naturally. The capability contributes to brain–computer interfaces, neurological disease diagnosis, and the understanding of complex brain functions.

Robotics Some robots are already being deployed in the healthcare industry. These include assisted laparoscopic surgical units and equipment. Improvements in haptic technology can increase the effectiveness and safety of these robots by providing doctors using them to remotely operate on patients with the tactile sensation of actually holding surgical tools. Robotics will also be increasingly useful to support aging populations. Assistive robots could help people move around, while other robots can help nursing and homecare workers provide essential functions such as bathing or cleaning.

Semiconductors Semiconductor chips are ubiquitous in modern medical equipment. Imaging devices such as magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound use embedded computers to generate images from electromagnetic radiation and sound waves penetrating or emanating from the human body.

Space The potential for space manufacturing can improve development of specialized pharmaceuticals, which can be made in a microgravity environment with minimal contaminants.

Civil Society

Artificial intelligence Because AI models are trained on existing datasets, they are likely to encode any biases present in these datasets. This affects model-based outcomes and decision making. Many facial recognition algorithms are better at identifying lighter-skinned faces than darker-skinned ones, leading to discrimination against people with the latter. Indiscriminate data collection can violate privacy and copyrights. Deepfakes used for misinformation and disinformation have personal, legal, and political impacts. The long-term nature and extent of AI's impact on employment—in terms of displacing some jobs and improving productivity in others—are still unknown.

Biotechnology and synthetic biology Different religious traditions may have different stances toward life or living systems, as well as different opinions as to whether the engineering of new life-forms violates any of their basic precepts. Another deliberation will be over who should have access to the benefits from synthetic biology given the risks to human and environmental safety from both malicious and unintentional acts.

Cryptography and computer security The nature of cryptography and encrypted communications raises questions about exceptional access regulations. These would require communications carriers and technology vendors to provide access to encrypted information to law enforcement agents or other bodies under specific legal conditions. Such information would be shared on the basis that encryption technology is also accessible to criminals and other malefactors. Opponents of exceptional access argue that implementing this capability weakens the security provided by encryption. Its supporters argue that the reduction in personal encryption security is worth the benefits of law enforcement's increased ability to catch and prosecute bad actors.

Energy technologies Continued creation of sustainable energy infrastructure requires new acquisitions of land to build generating stations and storage facilities. These can displace residents from private property and impact local property values, encouraging some to adopt a position of supporting windmills but "not in my backyard." The construction of nuclear power plants and facilities for storing radioactive waste is often met with opposition from those concerned about exposure to radiation in the environment.

Materials science There are many uncertainties about the long-term dangers and health concerns of nanoparticles released into the environment. Given this, there are important questions about how and to what extent regulations should be adopted to mitigate the risks potentially accompanying such releases. To resolve these questions, a consensus must be reached on the magnitude and severity of these risks and on appropriate remedies.

Neuroscience Neuroscience development is influenced by existing legal frameworks. The Controlled Substances Act, for instance, limits medical research on some substances that may have therapeutic effects. Meanwhile, cognitive and behavioral neuroscience have broad implications for public policy: For example, a basic aspect of criminal law is the nature and extent of an individual's responsibility for a criminal act. Currently, minors under eighteen years of age cannot be subject to the death penalty for crimes they have committed because adolescent brains are not considered fully developed; this puts minors at higher risk of impulsive, irrational thoughts and behaviors. As neuroscience advances, it could find evidence that reinforces or contradicts this principle and others.

Quantum technologies As quantum computers evolve, they may become capable of breaking the public-key encryption algorithms that protect data. Governments may have first access to such computers, but others may use them thereafter. Sensitive information belonging to ordinary citizens is already in the hands of malefactors, though in encrypted form, currently protected by today's public-key encryption algorithms. However, should quantum computers become able to break public-key encryption algorithms, this data will no longer be protected. Because that information is already in the hands of people who could exploit it in unencrypted form, there would be no recourse against their doing so.

Robotics Greater adoption of robotics will require moving workers to new roles as well as setting standards for human safety around robots. As robots assume more tasks, human workers will need education and training programs to undertake new roles and to benefit from robotics. Standards will also be needed to clarify limits to robotic applications. Ethical considerations warranting policy development include how to ensure data acquisition for training robots respects privacy and inclusiveness and how to set safety standards (e.g., Should the requirement be that a robot's performance is comparable to an average human's, or should it be near perfect?). Safety considerations for human-robot interactions will be an ongoing challenge.

Semiconductors Student interest in hardware design has dropped precipitously in favor of software-oriented jobs. Some estimates suggest that, given the current rates at which students with relevant degrees are graduating in the United States, 60 to 80 percent of jobs in semiconductor manufacturing will be unfilled by 2030.

Space In space, the rapid expansion of commercial assets and applications is raising important new policy considerations not covered by current norms. The increasing dependence of government on the private sector to provide space-based capabilities—including launch, vehicles, and space-based communications and internet access—vital to national security and economic growth raises questions about how to align public and private interests. Attempts at improvement have often stagnated due to nations' differing geopolitical aims. Dual-use space technologies and the challenge of getting private and government actors to cooperate will complicate crisis response.

CONCLUSION

This new edition of the *Stanford Emerging Technology Review* (*SETR*) has spotlighted ten pivotal technological domains that are shaping the future of science and innovation. Our extensive consultations with leading Stanford academics across scientific disciplines make clear that the coming decade will witness an unprecedented convergence of multiple fields that will drive technological change. Artificial intelligence (AI), fueled by increased computing power and more data, has potential to enhance human productivity dramatically and facilitate advancements across scientific fields, from drug discovery to breakthroughs in new materials. Synthetic biology and biotechnology promise groundbreaking applications in agriculture, healthcare, and industrial production.

Technological progress spans from the vast expanses of space to the microscopic world of nanoparticles. While technology itself is neither good nor bad, it's crucial for decision makers to comprehend the reach of technological change, its potential to either improve or disrupt societal norms—and the imperative for American leadership in navigating these expanding frontiers.

For decades, the prevailing approach to US science and technology policy has been to fund research at academic institutions and national laboratories, anticipate breakthroughs, and hope for positive outcomes. On the path toward the decisive establishment of US leadership in science and technology, this approach has served the nation well. However, the new technological and geopolitical landscape has presented the nation with a pivotal moment in which the research community can reinvent and reenergize itself so that it better supports innovation and serves the long-term interests of the American people.

A new strategy will acknowledge the responsibility of the research community—and especially of universities—to advance deep, thoughtful fundamental research, rebuild meaningful civic dialogue, and restore the confidence of the American public. Superficial window dressing is insufficient; going back to the way things were will not happen. At the same time, a new strategy must preserve what has been good for the United States: support for the research community that enables it to generate the breakthrough discoveries of tomorrow on which myriad technological advancements can be built. The advances we benefit from today stem from decades-old investments, yet the commitment to nurturing the future has weakened, compared to the past.

Ultimately, humans develop and use technology, and effective governance to maximize benefits and mitigate risks requires human guidance. Policymakers can establish frameworks that encourage innovation,

Ultimately, humans develop and use technology, and effective governance to maximize benefits and mitigate risks requires human guidance.

set priorities and strategies, align economic policies to foster innovation and maintain leadership, and bolster America's position in international competition.

As well as offering a look at individual emerging technologies, this publication also highlights common themes that emerge across them related to the development of science and technology. The importance of universities in the American innovation trifecta—government, academia, and industry—stands out as a crucial factor. As we noted at the start of this report, the US government is the only funder capable of making the large, sustained, and sometimes risky investments in the basic science conducted at universities that will be essential for future applications. Such support is going to be even more critical in the years ahead as other nations step up their own investments in fundamental research.

Maintaining a technological lead in a domain is distinct from gaining it. Engaging with expertise around the world, leveraging the potential of highly skilled immigrants, and sustaining robust domestic development of scientific expertise are essential to reinforcing American leadership in an increasingly competitive global landscape. Recognizing the evolving role of government in technology development is also vital. Innovations are no longer created and protected solely by state-backed research groups; private corporations and even individuals are developing more and more transformative technologies.

This paradigm shift is most evident in fields like AI and space exploration, where private companies are spearheading the creation of large language model systems and deploying innovative, highly advanced assets into space—a domain previously dominated by governments. The concentration of power in different hands has significant implications for technology access, priorities, and policy.

This edition of *SETR* started by asking the question, "What do policymakers need to know about emerging technologies from Stanford?" It serves as an initial step in providing the necessary and rapidly changing knowledge about these crucial technologies, their key takeaways, future implications, and potential policy concerns. The goal is to foster meaningful and ongoing discussions that can lead to effective and timely policymaking, even as technologies continue to evolve. We hope you found it useful, and we welcome feedback on how to make the publication even more impactful in the future—send your thoughts to SETReview2026@stanford.edu.

LEADERSHIP

Co-chairs

Condoleezza Rice is the Tad and Dianne Taube Director of the Hoover Institution and a senior fellow on public policy. She is the Denning Professor in Global Business and the Economy at the Stanford Graduate School of Business. She served as the sixty-sixth secretary of state (2005–9) and as the nineteenth national security advisor (2001–5). Rice served as Stanford University's provost from 1993 to 1999. She is author and coauthor of nine books. She received her PhD in political science from the University of Denver.

Jennifer Widom is the Frederick Emmons Terman Dean of the School of Engineering and the Fletcher Jones Professor in Computer Science and Electrical Engineering at Stanford University. Her research interests span many aspects of nontraditional data management. She is a fellow of the Association for Computing Machinery and a member of the National Academy of Engineering and of the American Academy of Arts and Sciences. She received her PhD in computer science from Cornell University.

Amy Zegart is the Morris Arnold and Nona Jean Cox Senior Fellow at the Hoover Institution and professor, by courtesy, of political science at Stanford University. She is also a senior fellow at Stanford's Institute for Human-Centered Artificial Intelligence and at its Freeman Spogli Institute for International Studies. The author of five books, she specializes in US intelligence, emerging technologies and national security, grand strategy, and global political risk management. She received her PhD in political science from Stanford University.

Director and Editor in Chief

Herbert S. Lin is senior research scholar at the Center for International Security and Cooperation and research fellow at the Hoover Institution, both at Stanford University. His research interests relate broadly to emerging technologies and national security. He is an elected fellow of the American Association for the Advancement of Science. He received his doctorate in physics from the Massachusetts Institute of Technology.

Managing Editor

Martin Giles is a policy fellow at the Hoover Institution and the executive director of Hoover's Technology Policy Accelerator at Stanford University. He is a seasoned executive, with editing and publishing experience at *The Economist* and *MIT Technology Review*, where he focused on developments in Silicon Valley and on emerging technologies. Giles was also the first editorial director of In-Q-Tel. He received his MBA from the University of Chicago Booth School of Business.

Faculty Council

Zhenan Bao is the K. K. Lee Professor in Chemical Engineering and professor, by courtesy, of chemistry and of materials science and engineering at Stanford University. She is a member of the National Academy of Sciences, the National Academy of Engineering, the American Academy of Arts and Sciences, and the National Academy of Inventors. Bao is known for her work on artificial electronic skin and NeuroString, which are enabling a new generation of skin-like electronics for regaining sense of touch for neuroprosthetics, human-friendly robots, human–machine interfaces, and seamless health-monitoring devices. She received her PhD in chemistry from the University of Chicago.

Dan Boneh is a professor of cryptography and electrical engineering at Stanford University, codirector of the Stanford Computer Security Lab, and a senior fellow at the Freeman Spogli Institute for International Studies. His research focuses on applied cryptography and computer security. He has authored over two hundred publications and is a member of the National Academy of Engineering. He received his PhD in computer science from Princeton University.

Simone D'Amico is an associate professor of aeronautics and astronautics and, by courtesy, of geophysics at Stanford University, where he serves as the W. M. Keck Faculty Scholar of Engineering. He is the founding director of Stanford's Space Rendezvous Laboratory and its Center for AEroSpace Autonomy Research (CAESAR) and is a science fellow at the Hoover Institution. His research explores the intersection of advanced astrodynamics, spacecraft navigation and control, and machine learning to enable future distributed space systems. He currently leads four satellite swarm and formation-flying projects for NASA and the National Science Foundation. He received his PhD in aerospace engineering from Delft University of Technology.

Drew Endy is the Martin Family University Faculty Fellow in Undergraduate Education (bioengineering); faculty codirector of degree programs for the Hasso Plattner Institute of Design (the d.school); core faculty at the Center for International Security and Cooperation; senior fellow, by courtesy, of the Freeman Spogli Institute for International Studies; and senior fellow, by courtesy, and science fellow at the Hoover Institution at Stanford University. He serves as president and director of the Biobricks Foundation and as director of the iGEM Foundation and the Biobuilder Educational Foundation. His research focuses on the foundations of synthetic biology along with broader societal aspects. He received his PhD in biotechnology and biochemical engineering from Dartmouth College.

Mark A. Horowitz is the Fortinet Founders Chair of the Department of Electrical Engineering and the Yahoo! Founders Professor in the School of Engineering at Stanford University. His research has contributed to early RISC (reduced instruction set computer) microprocessors, multiprocessor designs, and high-speed interfaces, and he currently works to create new agile design methodologies for analog and digital VLSI (very-large-scale integration) circuits. He received his PhD in electrical engineering from Stanford University.

Steven E. Koonin is Edward Teller Senior Fellow at Stanford University's Hoover Institution. Before joining Stanford in 2024, he was a professor at New York University, with appointments in the Stern School of Business, the Tandon School of Engineering, and the Department of Physics. Koonin's work focuses on climate science and energy technologies. He served as undersecretary for science at the US Department of Energy (2009–11) and prior to that was BP's chief scientist. He spent nearly thirty years at the California Institute of Technology as professor of theoretical physics and as provost. Koonin is a member of the National Academy of Sciences and the author of the book *Unsettled*, on climate science.

Fei-Fei Li is the inaugural Sequoia Professor in the Computer Science Department at Stanford University and founding codirector of Stanford's Institute for Human-Centered Artificial Intelligence. Her current research focuses on cognitively inspired artificial intelligence (AI), machine learning, deep learning, computer vision, robotic learning, and AI and healthcare. She received her PhD in electrical engineering from the California Institute of Technology and holds an honorary doctorate from Harvey Mudd College.

Allison Okamura is the Richard M. Weiland Professor of mechanical engineering in the School of Engineering and professor, by courtesy, of computer science at Stanford University. She is an executive committee member of the Stanford Robotics Center; affiliated faculty at Stanford Bio-X and Stanford's Institute for Human-Centered Artificial Intelligence; a senior fellow, by courtesy, and science fellow at the Hoover Institution;

and a fellow of the Institute of Electrical and Electronics Engineers. Her research interests include haptics, teleoperation, mixed reality, and medical and soft robotics. She received her PhD in mechanical engineering from Stanford University.

Kang Shen is the Frank Lee and Carol Hall Professor of biology and professor of pathology at Stanford University, where he serves as the Vincent V. C. Woo Director of the Wu Tsai Neurosciences Institute and as affiliated faculty at Stanford Bio-X. His research focuses on neuronal cell biology and developmental neuroscience. He has authored or coauthored more than one hundred journal articles. He received his PhD in cell biology from Duke University.

Jon Simon is the Joan Reinhart Professor and professor of applied physics in the Department of Physics at Stanford University. His research lies at the intersection of atomic, molecular, and optical physics, where he investigates quantum and classical matter composed of light. Simon's work explores how engineered photonic systems can emulate and reveal the complex behaviors of condensed-matter systems and how to leverage this unique control of photons to develop new hardware to control and read out quantum computers. He received his BS in physics from the California Institute of Technology and his PhD from Harvard University.

Advisory Board

Steven Chu is the William R. Kenan Jr. Professor of physics and professor of molecular and cellular physiology and of energy science and engineering at Stanford University. He previously served as US secretary of energy (2009–13) and as director of the Lawrence Berkeley National Laboratory. He received his PhD in physics from the University of California–Berkeley.

Robert Gates served as the twenty-second secretary of defense (2006–11) and as the director of central intelligence (1991–93) following nearly twenty-seven years in the CIA and on the National Security Council. He is the current chancellor of William & Mary and previously served as president of Texas A&M University. He received his PhD in Russian and Soviet history from Georgetown University.

Susan M. (Sue) Gordon is the former principal deputy director of national intelligence at the Office of the Director of National Intelligence (2017–19). She was formerly deputy director of the National Geospatial-Intelligence Agency (2015–17). She served twenty-seven years in the CIA, including as director of the CIA's Information Operations Center, and is now a leading voice on intelligence, cyber, space, and disruptive technologies as an independent director, advisor, and university fellow. She continues to shape national security as a member of the Council on Foreign Relations and the Defense Innovation Board.

John L. Hennessy is a professor of electrical engineering and computer science at Stanford University and the chairman of Alphabet Inc. He previously served as the tenth president of Stanford University (2000–16) and is the founder and director of the Knight-Hennessy Scholars Program, the world's largest fully endowed, university-wide graduate fellowship. He received his PhD in computer science from the State University of New York at Stony Brook.

Norbert Holtkamp is a science fellow at the Hoover Institution and professor of particle physics and astrophysics and of photon science at the SLAC National Accelerator Laboratory. He previously served as SLAC's deputy laboratory director (2014–22). He received his PhD in physics from the Technical University of Darmstadt in Germany.

Jerry McNerney is the former US representative for California's ninth congressional district (2007–23), where he served on the Committee on Energy and Commerce; the Veterans' Affairs Committee; and the Space, Science, and Technology Committee. He was elected to the California state senate in 2024 and is a senior policy advisor at Pillsbury Winthrop Shaw Pittman LLP. He received his PhD in mathematics from the University of New Mexico–Albuquerque.

Mary Meeker is a cofounder and general partner of BOND and serves on the boards of Block/Square, Genies, and Plaid. Her investments at BOND (and, previously, Kleiner Perkins) include Spotify, Waze, DocuSign, Ring, Checkr, Ironclad, and On Running. She was previously a managing director at Morgan

Stanley focused on emerging technology companies. She is also the coauthor of *USA Inc.: A Basic Summary of America's Financial Statements* and publisher of the widely distributed Internet Trends Report.

Lloyd B. Minor is the Carl and Elizabeth Naumann Dean of the School of Medicine and vice president for medical affairs at Stanford University. He is a professor of otolaryngology–head and neck surgery and professor, by courtesy, of bioengineering and neurobiology at Stanford University. Prior to joining Stanford, he was provost and senior vice president for academic affairs at Johns Hopkins University. He is an elected member of the National Academy of Medicine.

Peter L. Scher is the vice chairman of JPMorgan Chase & Co. He previously served in senior government positions, including as US special trade ambassador under President Clinton and as staff director of the US Senate Committee on Environment and Public Works. Prior to joining JPMorgan Chase in 2008, he was the managing partner of the Washington, DC, office of the law firm Mayer Brown LLP.

Eric Schmidt is cofounder, with his wife, Wendy, of Schmidt Sciences, a nonprofit organization working to advance science and technology to address global challenges. He currently chairs the Special Competitive Studies Project, a nonprofit initiative focused on strengthening America's long-term artificial intelligence and technological competitiveness in national security, the economy, and society. Previously, he was CEO and chairman of Google (2001–11) and executive chairman and technical advisor to Alphabet Inc. (2015–20). He received his PhD in electrical engineering and computer science from the University of California–Berkeley.

Thomas M. Siebel is the chairman and CEO of C3 AI and was the founder and CEO of Siebel Systems. He is a member of the colleges of engineering boards at the University of Illinois Urbana–Champaign and the University of California–Berkeley and is an elected member of the American Academy of Arts and Sciences.

Senior Research Staff

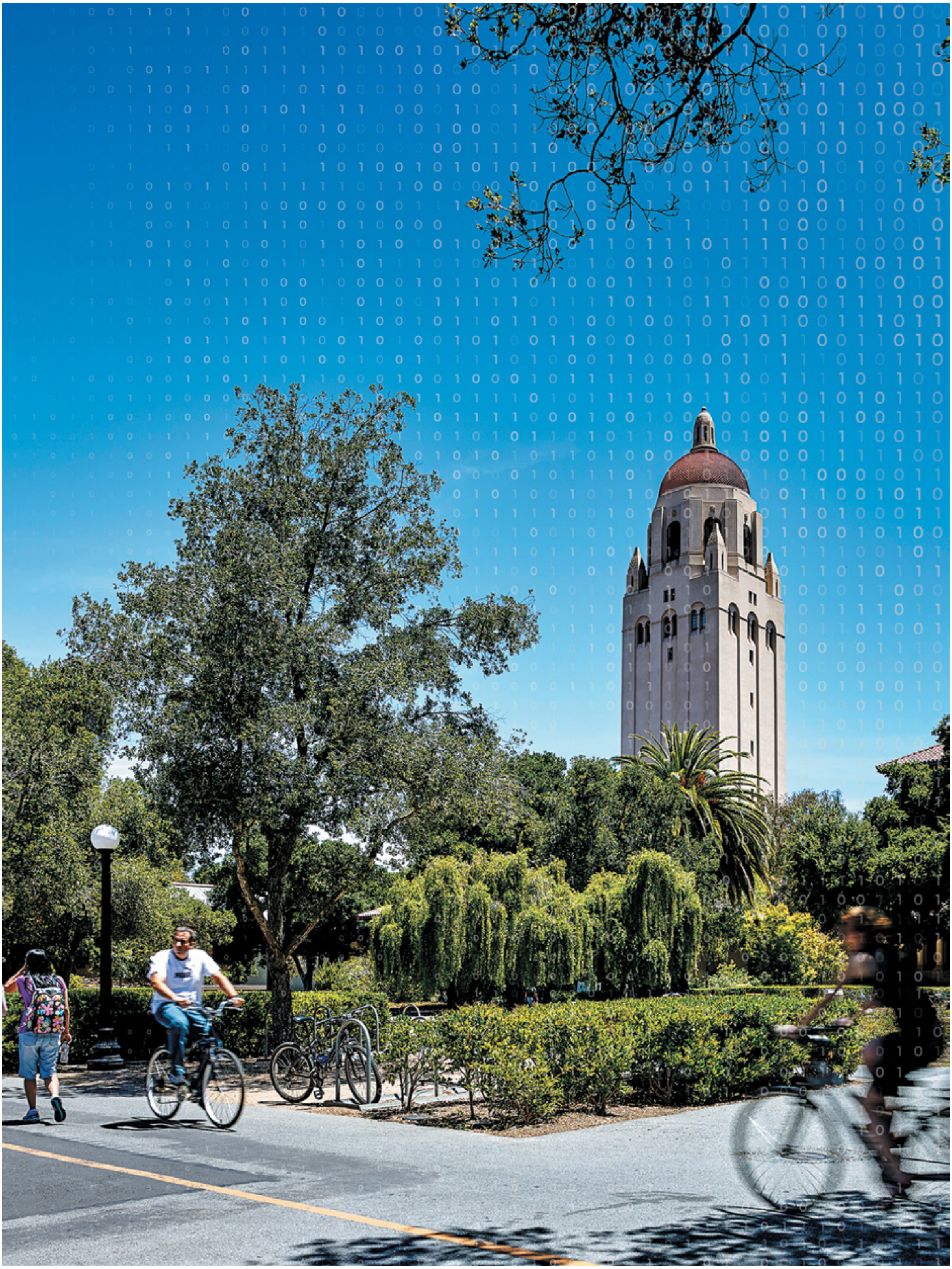
Emerson V. Johnston is a senior research assistant at the Hoover Institution's Technology Policy Accelerator at Stanford University. She is on leave from Stanford's master's program in history, where she specializes in science, technology, the environment, and medicine, and is a recent alumnus of Stanford's master's program in international policy, where she specialized in cyberpolicy and security.

ACKNOWLEDGMENTS

The members of the Stanford Emerging Technology Review (SETR) Faculty Council gave seminars on leading-edge work going on in their respective fields, and many of their faculty colleagues offered expert commentary and insights at these seminars. SETR faculty and fellows both contributed much effort in crafting entire chapters based on their deep understanding of their respective domains, including original ideas drawn from their own research. The SETR Advisory Board gave invaluable advice on how to interface with policymakers and the policy process. SETR staff and undergraduate research assistants did heroic work fact-checking and tracking down references, among many valuable tasks they performed. The Hoover Press did a wonderful job turning the manuscript into a final report.

For all of this, we are grateful.

Condoleezza Rice, Co-chair
Jennifer Widom, Co-chair
Amy Zegart, Co-chair
Herbert S. Lin, Director and Editor in Chief
Martin Giles, Managing Editor



This publication reflects updates through December 2025

32 31 30 29 28 27 26 7 6 5 4 3 2 1

Designer: Howie Severson
Typesetter: Maureen Forys

Image credits: Linda A. Cicero/Stanford News and iStock.com/PTC-KICKCAT92 (cover); iStock.com/mofuku (p. 22); iStock.com/wacomka (p. 38); iStock.com/FeelPic (p. 56); iStock.com/JONGHO SHIN (p. 70); iStock.com/Chartchai San-saneeyashewin (p. 88); iStock.com/ArtemisDiana (p. 102); iStock.com/PhonlamaiPhoto (p. 116); iStock.com/imaginima (p. 142); iStock.com/Floriana (p. 156); iStock.com/dima_zel (p. 170); Tim Griffith (p. 225)