STANFORD UNIVERSITY

THE STANFORD EMERGING TECHNOLOGY REVIEW 2023

A Report on Ten Key Technologies and Their Policy Implications

CHAIRED BY Condoleezza Rice, John B. Taylor, Jennifer Widom, and Amy Zegart

DIRECTED BY Herbert S. Lin

THE STANFORD EMERGING TECHNOLOGY REVIEW 2023

A Report on Ten Key Technologies and Their Policy Implications

CHAIRED BY
Condoleezza Rice
John B. Taylor
Jennifer Widom
Amy Zegart

DIRECTED BY
Herbert S. Lin

Stanford University
Stanford, California

TABLE OF CONTENTS

FOREWORD

Emerging technologies are transforming societies, economies, and geopolitics. This moment brings unparalleled promise and novel risks. In every era, technological advances buoy nations that develop and scale them—helping to save lives, win wars, foster greater prosperity, and advance the human condition. At the same time, history is filled with examples where slow-moving governments stifled innovation in ways policymakers never intended, and nefarious actors used technological advances in ways that inventors never imagined. Technology is a tool. It is not inherently good or bad. But its use can amplify human talent or degrade it, uplift societies or repress them, solve vexing challenges or exacerbate them. These effects are sometimes deliberate but often accidental.

The stakes of technological developments today are especially high. Artificial intelligence (AI) is already revolutionizing industries, from music to medicine to the military, and its impact has been likened to the invention of electricity. Yet AI is just one among many technologies that are ushering in profound change. Fields like synthetic biology, materials science, and neuroscience hold potential to vastly improve health care, environmental sustainability, economic growth, and more. We have experienced moments of major technological change before. But we have never experienced the convergence of so many technologies with the potential to change so much, so fast.

The *Stanford Emerging Technology Review* (*SETR*) is the first product of a major new Stanford technology education initiative for policymakers. Our goal is to help both the public and private sectors better understand the technologies poised to transform our world so that the United States can seize opportunities, mitigate risks, and ensure that the American innovation ecosystem continues to thrive.

Our efforts are guided by four observations:

1. **Policymakers need better resources to help them understand technological developments faster, continuously, and more easily**

Technology policy increasingly requires a more sophisticated understanding across a broad range of fields and sectors. Indeed, policymakers today include an expanding array of decision makers, from legislators and executive branch officials in Washington to state and local governments, investors, and corporate leaders. Too often, government leaders lack technical expertise to understand scientific developments, while technologists lack the policy expertise to consider and build security, safety, and other societal considerations into their products by design. Key takeaways of this report, for example, include the following findings that may be surprising and even counterintuitive to nonexperts:

- Artificial intelligence has received a great deal of media attention, but biotechnology could ultimately be as transformational to society as computing.
- Space technologies are increasingly critical to everyday life, from GPS navigation to banking. But space is a planetary resource that is rapidly becoming congested and contested—with thousands of new commercial satellites and an estimated million pieces of space debris that could threaten access to these global commons.
- The most significant challenge to achieving sustainable energy is scale; simply providing a 72-hour supply of backup energy worldwide would take two hundred years of lithium-ion battery production.

- Cryptocurrencies are not the most important issue in cryptography today, and they are not synonymous with blockchain, which has widespread applications.

As these examples suggest, policymakers need better, easy-access resources to help them understand technological basics and new discoveries before crises emerge; to focus their attention on the most important issues; to better assess the policy implications; and to see over the horizon to shape, accelerate, and guide future technological innovation and applications. We need a new model of technology education for nontechnical leaders. This report aims to be a first, important step.

2. 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.

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, for example, American and Soviet nuclear scientists and policymakers worked together to reduce the risk of accidental nuclear war through arms control agreements and safety measures. Today, China's rise poses many new challenges. Yet maintaining a robust global ecosystem of scientific cooperation remains essential—and it does not happen by magic. It takes work, leadership, and a fundamental commitment to freedom to sustain the openness essential for scientific discovery. Freedom 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 or persecution. In short, it matters whether the innovation ecosystem is led by democracies or autocracies. The United States has its flaws and challenges, but this country remains the best guarantor of scientific freedom in the world.

3. Academia's role in American innovation is essential yet increasingly at risk

The US innovation ecosystem has three pillars: the government, the private sector, and the academy. Success requires that all three remain robust and

> **The United States has its flaws and challenges, but this country remains the best guarantor of scientific freedom in the world.**

actively engaged. Throughout history, America's research universities have generated transformational scientific discoveries, from the invention of the polio vaccine to rocket fuel. Universities have also been the seedbeds of policy innovations, from nuclear deterrence theory to behavioral economics. And they have played a vital role in training the next generation.

Today, however, innovations are increasingly emerging from the private sector, often alongside academia. The funding sources for innovation have shifted, too—in deeply worrying ways. The US government is the only funder capable of making large and risky investments in the basic science conducted at universities (and national laboratories) that is essential for future applications. Yet federal research and development (R&D) funding has plummeted since the 1960s, from 1.86 percent of GDP in 1964 to just 0.66 percent of GDP in 2016.[1] Although private sector investment in technology companies and associated university research has increased substantially, it is no substitute; federal funding of university research leads universities to study different technological challenges and opportunities than industry funding does. As a Council on Foreign Relations innovation task force report concluded:

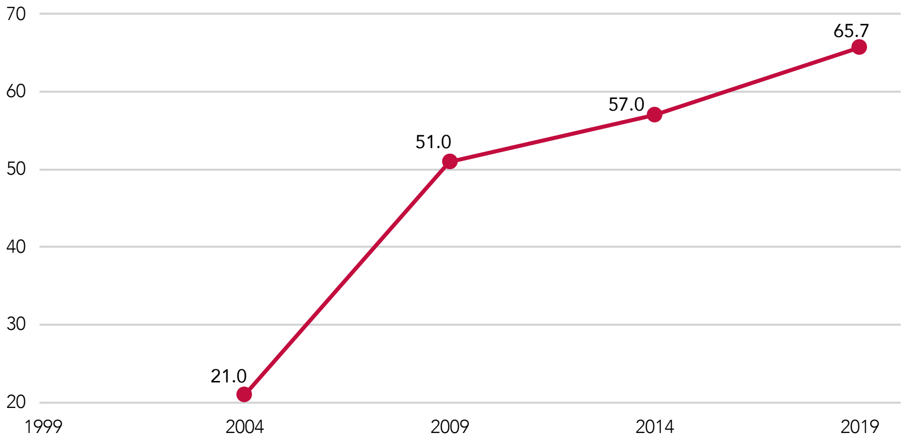
> U.S. leadership in science and technology is at risk because of a decades-long stagnation in federal support and funding for research and development. Private-sector investment has risen, but it is not a substitute for federally funded R&D directed at national economic, strategic, and social concerns.[2]

To be sure, the rising influence 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. Universities and companies are not the same. Companies must answer to investors and shareholders who expect returns on their capital investments, so they tend to focus on technologies that can be commercialized in the foreseeable future. Research universities, by contrast, operate on much longer time horizons without regard for profit, engaging in fundamental research at the frontiers of knowledge that has little if any foreseeable commercial benefit. This fundamental research is the foundation for future applications that may take years, even decades, to emerge. The "overnight success" of the COVID mRNA vaccine in 2021, for example, was the result of thirty years of university research. Similarly, it took decades of research in number theory—a branch of pure mathematics—to develop the modern cryptography that is widely used to protect data.

Today, 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 2020, two-thirds of students who received PhDs in artificial intelligence at US universities took industry jobs, leaving fewer faculty to teach the next generation (see figure F.1).[3] Only a handful of the world's largest companies have both the talent and the enormous compute power necessary for developing sophisticated large language models like GPT-4. No university comes close.

These trends have several concerning implications.[4] Among them: 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 large language models 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 future generations of the best AI minds will continue to flock there (see figure F.1)—hollowing out university faculty and eroding the nation's ability to conduct broad-ranging foundational research in the field.

FIGURE F.1 Percentage of AI PhDs hired by industry



Source: Nur Ahmed, Muntasir Wahed, and Neil C. Thompson, "The Growing Influence of Industry in AI Research," *Science* 379, no. 6635 (March 2023): 884–86.

4. 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; the university lies at the heart of the innovation ecosystem. Stanford faculty, researchers, and former students have founded Alphabet, Cisco Systems, Hewlett-Packard, 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, laboratories, and kitchens. 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 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 individuals 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. But in this moment of rapid technological change, we must do more. We are delighted to launch this unprecedented collaboration between Stanford's Hoover Institution, the School of Engineering, and the Institute for Human-Centered Artificial Intelligence to bring policy analysis, social science, science, medicine, and engineering together.

The *Stanford Emerging Technology Review* originated from conversations we had last year with senior US government officials who came to campus and asked, "What do we need to know about emerging technologies at Stanford?" No one person had a good answer, so we convened leading scholars across fields for briefings. The impact of that day was powerful and revealing: it was a one-off event, and it was not enough. We also discovered that many of our leading faculty in different science and engineering fields did not know one another. Together we realized that although Stanford is one of the world's leading research universities, we did not know what we knew. And fragmentation was hindering our policy impact.

So we founded the Stanford Emerging Technology Review (SETR), an enduring initiative to harness the latest insights from leading scholars in ten of the most important fields today, bring these scholars together to share their research with colleagues across disciplines, and work collaboratively to enhance policy education and impact for the nation.

We selected these ten areas as a starting point, not an end point. We wanted to begin by leveraging areas of deep expertise at Stanford and covering technologies widely recognized as essential for expanding American economic prosperity, advancing democratic values, and protecting the security of the nation. But science is always moving, and we expect that future reports may focus on different areas or divide fields in different 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 ten 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-Stop Shopping but Not a One-Time Product

This report is intended to be a useful "one-stop shopping" primer that covers ten key emerging technology areas: artificial intelligence, biotechnology and synthetic biology, cryptography, materials science, neuroscience, nuclear technologies, robotics, semiconductors, space technologies, and sustainable energy technologies. While this is nowhere near an exhaustive list of technology research areas at Stanford, these ten fields are rapidly shaping American society today and promise to gain importance in the coming years. Our reviews of each technology field were led by world-renowned Stanford tenured faculty members who also delivered lectures covering their fields in SETR seminars (their bios can be found in the Contributors section on page 151). The SETR team also included eighteen postdoctoral scholars and eleven undergraduate research assistants who spent the last year interviewing leading faculty across Stanford in different subfields, conducting research, and drafting background materials. Overall, they conducted seventy-five interviews spanning faculty from thirty departments on the key developments, barriers, bottlenecks, needs, opportunities, and implications in their respective fields.

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 key developments and advances in the field. Finally, each chapter concludes by offering an "over-the-horizon" outlook that covers crucial considerations for policymakers over the next few years. The report ends with two chapters that look across the ten technologies, offering analysis of common trends, key differences, and implications for economic growth, national security, environmental and energy sustainability, human health, and civil society.

Three points bear noting. **First, we offer no specific policy recommendations.** That is by design. Washington is littered with reports offering policy recommendations that were long forgotten, overtaken by events, or both. We want to provide a reference resource that endures—a report that is updated and issued annually, a guide that can inform successive generations of policymakers about evolving technological fields and their implications.

Second, *SETR* offers *a* view from Stanford, not *the* view from Stanford. There is no single view of anything in a university. Individual faculty members involved in this report may not agree with everything in it. Other members of their departments would probably offer a different lay of the technology landscape with varying assessments about important developments and over-the-horizon issues. The report is intended to reflect the best collective judgment about the state of these ten fields—guided by leading experts in those fields.

Third, this report is just the beginning. In the months ahead, SETR will be producing additional articles and reports, holding briefings in California and Washington, DC, and launching multimedia educational products. Our goal is ambitious: developing a new model to help policymakers understand tech issues in a more real-time, continuous, rigorous, and user-friendly way.

Ensuring American leadership in science and technology requires all of us—academia, industry, government—to keep listening, learning, and working together. We hope the *Stanford Emerging Technology Review* starts meaningful and lasting conversations about how an innovation ecosystem benefits us all. The promise of emerging technology is boundless if we have the foresight to understand it and the fortitude to embrace the challenges.

Condoleezza Rice
John B. Taylor
Jennifer Widom
Amy Zegart

Co-chairs, Stanford Emerging Technology Review

NOTES

1. Council on Foreign Relations, *Innovation and National Security: Keeping Our Edge*, Independent Task Force Report No. 77, James Manyika and William H. McRaven, chairs, 2019, 10, https://www.cfr.org/keeping-our-edge/pdf/TFR_Innovation_Strategy.pdf.

2. Council on Foreign Relations, *Innovation and National Security*, 21.

3. Nur Ahmed, Muntasir Wahed, and Neil C. Thompson, "The Growing Influence of Industry in AI Research," *Science* 379, no. 6635 (March 2023): 884–86, https://www.science.org/doi/abs/10.1126/science.ade2420.

4. 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?," 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 more easily understand the burgeoning and complex array of technological developments—more easily and more continuously.

The *Stanford Emerging Technology Review* 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 2023 Focus Technologies

Artificial Intelligence
Biotechnology and Synthetic Biology
Cryptography
Materials Science
Neuroscience
Nuclear Technologies
Robotics
Semiconductors
Space
Sustainable Energy Technologies

These particular fields were chosen for the 2023 report because they leverage areas of deep expertise at Stanford and cover many critical and emerging technologies identified last year by the Office of Science and Technology Policy in the White House and other US government departments. However, *SETR* focus technologies are likely to change over time, not because anyone "got it wrong," 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 standalone chapter that gives an overview of the field, highlights key developments, and offers an over-the-horizon view of important technological and 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 areas: economic growth, national security, environmental and energy sustainability, health and medicine, and civil society.

Three tensions run throughout and are worth keeping in mind.

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

2. Technical depth and breadth This report intentionally skews toward breadth, offering a 30,000-foot view of a vast technological landscape in one compendium. Readers should consider this report as an introductory course. *SETR* will issue deeper-dive reports and other educational tools in the months ahead that will offer more advanced examinations of each field.

3. 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—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.[1] 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 (AI)

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 last year, the main AI-related headline was the rise of large language models (LLMs) like GPT-4, on which the chatbot ChatGPT is based.

KEY CHAPTER TAKEAWAYS:

- AI is a foundational technology that is advancing other scientific fields and, like electricity and the internet, has the potential to transform how society operates.
- Even the most advanced AI has many failure modes that are unpredictable, not widely acknowledged, not easily fixed, not explainable, and capable of leading to unintended consequences.
- There is substantial debate among AI experts about whether AI poses a long-term existential risk to humans, and whether the most important risks are current weaknesses of AI.

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 it 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 burgeoning, contributing around 5 percent to US GDP with a historical doubling time of about seven years.
- Synthetic biology is third-generation biotechnology, complementing domestication and breeding (the first generation) and gene editing (the second generation).
- The United States is struggling to grasp the scale of the bio-opportunity, the strategic ramifications unique to network-enabled biotechnologies, and the possibilities and perils of distributed biomanufacturing.

Cryptography

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, including 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 will never be enough to secure cyberspace.
- Cryptocurrencies have received a great deal of media attention, but they are not the most important issue in cryptography today.
- Cryptocurrencies use blockchain technology, but they are not the same; blockchain has many other important and promising applications.

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. Materials science contributions have led to better semiconductors, "smart bandages" with integrated sensors and simulators that can accelerate healing, more easily recyclable plastics, more energy efficient and flexible solar cells, and stronger aircraft parts.

KEY CHAPTER TAKEAWAYS:

- Materials science is a foundational technology that underlies advances in many other fields, including robotics, space, energy, and synthetic biology.
- Materials science will exploit AI as another promising tool to predict new materials with new properties and identify novel uses for known materials.
- The structure of funding in materials science does not effectively enable transition from innovation to implementation. Materials-based technology that has been thoroughly tested at the bench scale may be too mature to qualify for basic research funding (because the high-level basic science is understood) but not mature enough to be directly commercialized by companies.

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:

- Popular interest in neuroscience vastly exceeds the actual current scientific understanding of the brain, giving rise to overhyped claims in the public domain that revolutionary advances are just around the corner.
- Advances in computing have led to progress in several areas, including understanding and treating addiction and neurodegenerative diseases, and designing brain-machine interfaces.
- American leadership is essential for establishing and upholding global norms about ethics and human subjects research in neuroscience.

Nuclear Technologies

Nuclear technologies involve producing energy with potential applications for electricity generation, medicine, and weapons. There are two major nuclear processes: (1) fission, which is the process of splitting the nucleus of a particular type of element; and (2) fusion, which produces energy by causing two atoms to collide and fuse together. Nuclear power plants have used controlled fission chain reactions for decades. In the past year, however, Lawrence Livermore National Laboratory achieved a milestone breakthrough, raising hopes that fusion might someday be controlled to drive electrical generators without the long-lasting radioactive waste that fission produces.

KEY CHAPTER TAKEAWAYS:

- Nuclear fission offers a promising carbon-free power source that is already in use but faces safety and proliferation concerns, economic obstacles, and significant policy challenges to address long-term radioactive waste disposal.
- Nuclear fusion recently achieved an important milestone by demonstrating energy gain in the laboratory for the first time. However, further research breakthroughs must be achieved in the coming decades before fusion can be technically viable as an energy alternative.
- Many believe that small modular reactors (SMRs) are the most promising way to proceed with nuclear power, but some nuclear experts have noted that SMRs do not solve the radioactive waste disposal problem.

Robotics

Robotics is an integrative field that draws on advances in multiple technologies rather than a single discipline. "What is a robot?" is a harder question 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:

- Although robots today are mostly used for the Three Ds (dull, dirty, or dangerous tasks), in the future they could be used for almost any task involving physical presence, because of recent advances in AI, decreasing costs of mobile component technologies (e.g., cameras in smartphones), and designs enabled by new materials and structures.
- Robotics has and will transform many industries through elimination, modification, or creation of jobs and functions.

- Understanding and communicating how robots will affect people's lives directly in their physical spaces (e.g., security robots in malls) as well as more existentially (e.g., transitioning jobs like truck driving from human-driven to autonomous vehicles) will shape how the United States accepts and benefits from robotic technologies.

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 actually manufacturing chips in large, specially designed factories called "fabs." Because fabs involve highly specialized equipment and facilities—the "clean rooms" in which chips are made require air that is one thousand times more particle-free than a hospital operating room—they are extremely expensive to build and require expertise to operate. US companies still play a leading role in semiconductor design, but US semiconductor-manufacturing capacity has plummeted and now lags dangerously behind Taiwan Semiconductor Manufacturing Company (TSMC) and Korea's Samsung. The Creating Helpful Incentives to Produce Semiconductors and Science Act of 2022 (CHIPS Act of 2022) was intended to help the US semiconductor industry regain a foothold in fabrication, but progress is expected to take years, if not decades. Because of the cost and complexity involved, success remains uncertain. At the same time, we are reaching the limits of exponential technical and cost improvements in the chip fabrication process, known as Moore's law. Until now, systems and software have been designed with the expectation that semiconductor capabilities would dramatically increase and costs would decrease over time. That is unlikely to be the case in the future, with profound implications for the development of hardware and software as well as the innovation that depends on it.

KEY CHAPTER TAKEAWAYS:

- Moore's law, which for fifty years has predicted rapid increases in semiconductor capabilities at decreasing costs, is now ending, raising profound implications for the future of hardware and software development.

- Recent research has identified methods that allow innovations in materials, devices, fabrication, and hardware to be added to existing process or systems at low incremental cost. These methods need to be further developed since they will be essential to continue to improve the computing infrastructure we all depend on.

- Quantum computing may solve certain specialized problems, but experts debate whether it can ever achieve the rapid, consistent, predictable performance growth that semiconductors have enjoyed.

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. Today, more than eight thousand working satellites circle the planet, many no larger than a loaf of bread. Some operate in constellations that can revisit the same location multiple times a day and offer image resolutions so sharp they can identify different car models driving on a road.

KEY CHAPTER TAKEAWAYS:

- Space technologies are increasingly critical to everyday life (e.g., GPS navigation, banking, missile defense, internet access, and remote sensing).
- Space is a finite planetary resource. Dramatic increases in satellites, debris, and competition are threatening access to this global commons.
- Private-sector actors play a critical and growing role in many aspects of space-based activities (e.g., launch, vehicles, and communications), because they offer better, cheaper, and rapidly deployable capabilities.

Sustainable Energy Technologies

This vital strategic resource for nations generally involves generation, transmission, and storage. In recent years it has also come to include carbon capture and carbon's removal from the atmosphere. Energy mix and innovation are key to efforts to address climate change.

KEY CHAPTER TAKEAWAYS:

- The most significant challenge to achieving sustainable energy is scale. Countries will need to source, manufacture, and deploy massive generation, transmission, and storage capabilities to meet global energy needs.
- Because global energy needs are vast, no single technology or breakthrough will be enough.
- Over-the-horizon challenges include decentralizing and modernizing the country's electricity grids and achieving greater national consensus about energy goals to enable strategic and effective R&D programs and funding.

Important Crosscutting Themes

Chapter 11 discusses twelve themes that cut across the technological areas. These are:

1. Different risks arise from moving too fast and moving too slowly. Innovation that emerges too fast threatens to disrupt the status quo around which many national, organizational, and personal interests have coalesced. It is also more likely to lead to unintended consequences and give short shrift to security, safety, ethics, and geopolitics. Innovation that moves too slowly increases the likelihood that a nation will lose the technical, economic, and national security advantages that often accrue to first movers in a field.

2. Ideas and human talent play a central role in scientific discovery and cannot be manufactured at will. They must be either domestically nurtured or imported from abroad. Today, both paths for generating ideas and human talent face serious and rising challenges.

3. The US government is no longer the primary driver of technological innovation or funder of research and development. Historically, technological advances (e.g., semiconductors, the internet, jet engines) were funded and advocated by the US government. Today, private sector R&D investment is playing a much larger role, raising important concerns about how to ensure that the national interest is well considered and that basic science—which is an important foundation for future innovation—remains strong.

4. There is a trend toward increasing access to new technologies worldwide. Even innovations that are US born are unlikely to remain in the exclusive control of American actors for long periods.

5. The synergies between different technologies are large and growing. Advances in one technology often support advances in other technologies.

6. The path from research to application is often not linear. Many believe that technological breakthroughs arise from a step-by-step linear progression where basic research leads to applied research, which then leads to development and prototyping and finally to a marketable product. Yet innovation often does not work this way. Many scientific developments enhance understanding but never advance to the marketplace. Many marketable products emerge in nonlinear fashion, after many rounds of feedback *between* phases. Other products emerge only when several different technologies acquire maturity.

7. Technological innovation occurs in both democracies and autocracies, but different regime types enjoy different advantages and challenges. Democracies provide greater freedom for exploration, while authoritarian regimes can direct sustained funding and focus on the technologies they believe are most important.

8. The speed of change is hard even for leading researchers to anticipate. Technology often progresses in fits and starts, with long periods of incremental results followed by sudden breakthroughs.

9. Nontechnical factors often determine whether new technologies succeed or fail. Adoption of new technologies hinges on economic viability and societal acceptability, not just scientific proof-of-concept and engineering feasibility.

10. US universities play a pivotal role in the innovation ecosystem that is increasingly at risk. Although the US government frequently talks about the importance of public-private partnerships in emerging technology, universities also play a pivotal and often underappreciated role. They are the only organizations with the mission of pursuing high-risk research that may not pay off commercially for a long time, if ever. That high-risk focus has yielded high-benefit payoffs in a wide range of fields.

11. Sustaining American innovation requires long-term government R&D. Investments with clear strategies and sustained priorities are crucial, not the increasingly common wild swings from year to year.

12. Cybersecurity is an enduring concern for every aspect of emerging technology research. State and nonstate actors will continue to threaten the confidentiality, integrity, and availability of information that is crucial for emerging technology research and development.

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.

NOTES

1. Andrew Pettegree and Arthur der Weduwen, *The Bookshop of the World: Making and Trading Books in the Dutch Golden Age* (New Haven, CT: Yale University Press, 2019), 70–72.

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 which has been the burden of the common man for ages past. . . . 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]

The importance of science and technology (S&T) remains essential to our national interests. Advances in S&T are closely tied to national needs in transportation, agriculture, communication, energy, education, environment, health, and defense—as well as 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, 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 sciences, and nanotechnology will impact both "hard power" issues—defense, arms control, nonproliferation—and "soft power" concerns, such as climate change, infectious and chronic diseases, energy supply and demand, and sustainable development.[3]

S&T is one important battleground for seeking advantage in geopolitical competition, as advances in S&T 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 leveraged alongside strong public policy if those advances 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 in the long run. Attempting to restrict S&T transfer to other nations may delay its spread, but the first successful demonstration of a technological advance on our part 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 S&T advances originating in one nation often benefit others. For example, the internet and GPS are US-born innovations whose uses have spread around the world—and the United States itself has gained from that spread.
- International competition does not occur only with adversaries. Our allies and partners also compete in the S&T space, developing technology or deploying policy that can leave the United States at a disadvantage.

Policy for Science and Technology

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; regulation; public procurement; funding and other aid to strategic sectors; and labor force training and education.

On the other hand, policy need not be directed at S&T to have a meaningful impact. 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 Science and Technology 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 and think tanks, as well as 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 the others, and the discussion below does so in alphabetical order. Indeed, one of the perhaps 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.

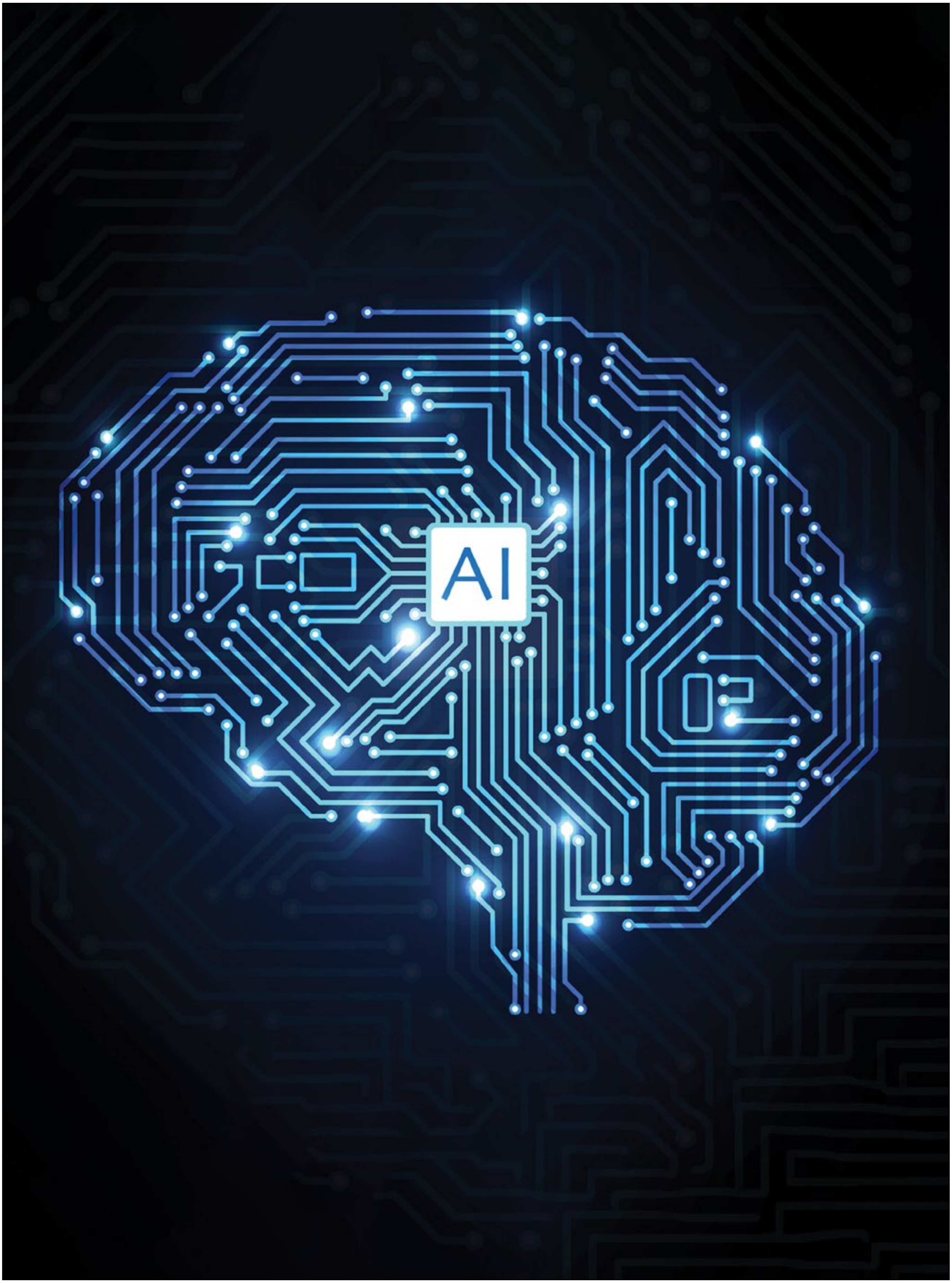
The description of each field is divided into three parts. The first part is an **overview** of the field. The second part addresses noteworthy **key developments** in the field that are relevant to understanding the field from a policy perspective. The last part, **over the horizon**, is itself subdivided into three sections: the potential impact of the field in the future (i.e., the field's potential over-the-horizon impact); the likely challenges facing innovation and implementation; and relevant policy, legal, and regulatory issues.

NOTES

1. Vannevar Bush, Science: *The Endless Frontier* (Washington, DC: 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* (Washington, DC: 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*, December 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*, December 2004, https://www.dni.gov/files/documents/Global%20Trends_Mapping%20the%20Global%20Future%202020%20Project.pdf.


AI

01

ARTIFICIAL INTELLIGENCE

KEY TAKEAWAYS

- AI is a foundational technology that is advancing other scientific fields and, like electricity and the internet, has the potential to transform how society operates.
- Even the most advanced AI has many failure modes that are unpredictable, not widely appreciated, not easily fixed, not explainable, and capable of leading to unintended consequences.
- There is substantial debate among AI experts about whether AI poses a long-term existential risk to humans, and whether the most important risks are current AI weaknesses.

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 on 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 are possible with human capabilities.

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 recent years. The science of computing, worldwide availability of networks, and civilization-scale data—all that collectively underlies the AI of today and tomorrow—promises 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 interact directly with sophisticated AI applications for a multitude of everyday activities.

The global AI market was worth $136.55 billion in 2022, with North America receiving 36.8 percent of total AI revenues.[2] A Stanford University study found that total private investment in artificial intelligence exceeded $93 billion in 2021, a twofold increase in capital from 2020.[3] While artificial intelligence start-ups received roughly 9 to 10 percent of global venture capital investment in recent years,[4] global AI start-up funding slowed considerably in 2022, dropping from an all-time high of roughly $18 billion in the third quarter of 2021 to approximately $8.3 billion in quarter three of the following year.[5] Generative AI, discussed below, is estimated to raise global GDP by $7 trillion and lift productivity growth by 1.5 percent over a ten-year period, if adopted widely.[6]

What subfields are considered part of AI is a matter of ongoing discussion, and the boundaries between these fields are often fluid. Some of the core subfields include:

- Computer vision, enabling machines to recognize and understand visual information from the world, converting it into digital data and making decisions based on it

- 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 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

Most of today's AI is based on machine learning, though it draws on other subfields as well. Machine learning requires data and computing power—often called compute[7]—and much of today's AI research requires access to these on an enormous scale.

Artificial intelligence requires large amounts of data from which it can learn. These data can take various forms, including text, images, videos, sensor readings, and more. The quality and quantity of data play a crucial role in determining the performance and capabilities of AI models. Without sufficient and high-quality data, AI models may generate inaccurate or biased outcomes. (Roughly speaking, a model is developed to solve a particular problem—different problems call for different models, and for problems that are sufficiently different from each other, entirely new models need to be developed.) Research continues today on how to train models incrementally, starting from an existing model and then using a much smaller amount of specially curated data to refine the performance of those models for specialized purposes.

For a sense of scale, one AI model recently in the news is GPT-4. Estimates of the data required to train this model suggest that around a million books (hundreds of gigabytes of text) were drawn from billions of web pages and scanned books. The hardware requirements for computing power are also substantial. The costs to compute the training of GPT-4, for example, were enormous. Reports indicate the training of this application took about twenty-five thousand Nvidia A100 GPU deep-learning chips—at a cost of $10,000 each—running for about one hundred days.[8] Doing the math and noting that other hardware components were likely also needed suggests the overall hardware costs for GPT-4 were at least a few hundred million dollars. And the chips underlying this hardware are specialty chips generally fabricated offshore.[9] (Chapter 8 on semiconductors discusses this point at greater length.)

Last, training AI models is an energy-intensive activity. One estimate of the electricity costs of training a large language model such as GPT-4 pegs the figure at about fifty million kilowatt-hours.[10] Then once it's up and running, the energy cost of a query on ChatGPT is around 0.002 of a kilowatt-hour.[11] Given hundreds of millions of queries per day, the operating energy cost of ChatGPT might be a few hundred thousand kilowatt-hours, at a cost of several tens of thousands of dollars, per day.

AI can automate a wide range of tasks. But AI also has particular promise in augmenting human capabilities and further enabling people to do what people are best at doing.[12] AI systems can work alongside people, complementing and assisting rather than replacing them. Some present-day examples include:

Health Care

- **Medical diagnostics** AI systems that can predict and detect the onset of strokes and subsequently perform automated triage, mobile viewing, and secure communication across several specialties and diseases qualified for Medicare reimbursement in 2020.[13]

- **Drug discovery** An AI-enabled search identified a compound that inhibits the growth of a bacterium responsible for many drug-resistant infections (e.g., pneumonia, meningitis) by sifting through a library of seven thousand potential drug compounds for an appropriate chemical structure.[14]

- **Robotic assistants** Mobile robots can carry out health care–related tasks such as making specialized deliveries, disinfecting hospital wards, and assisting physical therapists, thus supporting nurses and enabling them to spend more time with face-to-face human interactions.[15]

Agriculture

- **Production optimization** AI-enabled computer vision helps some salmon farmers sort fish into the right size to keep, thus off-loading the labor-intensive task of sorting fish.[16]

- **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 entire fields, in some cases reducing herbicide use by as much as 90 percent.[17]

Logistics and Transportation

- **Resource allocation** AI enables some commercial shipping companies to predict ship arrivals five days in the future with high accuracy, thus allowing real-time allocations of personnel and schedule adjustments.[18]

Law

- **Legal transcription** AI enables the real-time transcription of legal proceedings and client meetings with reasonably high accuracy, and some such services are free of charge.[19]

- **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—and even write case summaries.[20]

Key Developments

Large Language Models

Large language models (LLMs) are AI systems trained on very large volumes of written text to recognize, summarize, and 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. A simple example might be that the word sequence "thank you" is far more likely to occur than "thank zebras." The resulting systems, which include chatbots such as ChatGPT, Bard, and Claude, generate output surprisingly similar to that of humans across a wide range of subjects, including computer code, poetry, legal case summaries, and medical advice. LLMs are examples of foundation models, which are machine-learning models trained on big datasets that can drive a large number of applications.[21] LLMs are also an example of generative AI, a type of AI that can produce new content (e.g., text, images, sounds, animation) based on how it has been trained and the inputs it is given.

Computer Vision

In recent years, computer vision has made substantial progress on a number of important problems, including:

- Image classification (categorizing objects in images)
- Facial detection and recognition (finding faces in images and then matching those faces to existing face images)
- Medical image segmentation (identifying an organ in an image and isolating the portions of the image associated with that organ)
- Object recognition (identifying and localizing instances of objects in images)
- Activity recognition (identifying human activity depicted in a video, e.g., a human being sitting or walking)[22]

AI-Enabled Scientific Discovery

Over the past few years, AI models using large amounts of scientific data have been the accelerant of several scientific discoveries. Prominent examples include protein structure predictions for multiple proteins associated with SARS-CoV-2,[23] the use of AI models to discover new antibodies,[24] and the improvement of plasma control procedures for nuclear fusion.[25]

Existential Concerns about AI

Large language models have generated considerable attention because of their apparent sophistication. Indeed, their capabilities have led some to suggest that LLMs are the initial sparks of artificial general intelligence (AGI).[26] AGI is artificial intelligence capable of performing any intellectual task that a human can perform, including learning. But because it would run on a computer, it is likely to learn much faster than humans—outstripping human capabilities in short order.

The prospect that artificial general intelligence will soon be achieved has raised substantial debate about its risks. In May 2023, a number of senior and respected AI researchers released a statement saying that "mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks, such as pandemics and nuclear war." Concerned about the speed at which the power of AI-enabled systems is growing, they worry that in the absence of good governance, future systems could pose existential risks to humanity.

Others suggest that focusing on low-probability doomsday scenarios distracts from the real and immediate risks that AI poses today.[27] Instead, they argue, AI researchers should prioritize addressing the harms AI systems are already causing, like biased decision making and job displacement. These problems are the ones on which governments and regulators should be focusing their efforts.

A National Artificial Intelligence Research Resource

As it stands today, AI models such as GPT-4 can be developed only by large industrial actors with the resources to build and operate large data and compute centers—companies such as Google, Microsoft, and Meta (previously Facebook). Traditionally, academics and others in civil society have undertaken research to understand the potential societal ramifications of AI, but with large companies controlling access to these AI systems, they cannot do so independently.

For this reason, a bipartisan group of legislators in July 2023 proposed a bill to establish the National Artificial Intelligence Research Resource (NAIRR) as a shared national research infrastructure that provides civil society researchers greater access to the complex resources, data, and tools needed to support research on safe and trustworthy artificial intelligence. Even so, the scale of government resources proposed is a factor of five or ten lower than what the private sector is willing and able to invest.[28]

Over the Horizon

Impact of New Technologies

New technologies often have positive and negative impacts. Potential positive impacts of new AI technologies are most likely to be seen in the applications they enable for societal use. These may include:

- Truck drivers can off-load to AI the most boring and time-consuming aspects of their jobs—the long-haul drives—and still retain those aspects of their jobs that require human-centered interactions, usually involving the first and last miles of their routes.
- Smart AI sensors and cameras can improve patient safety in intensive care units, operating rooms, and even at home by improving health-care providers' and caretakers' ability to monitor and react to patient health developments, including falls and injuries.[29]

Potential negative AI impacts likely will emerge from known problems with the current state-of-the-art AI and from unfettered success with AI in the future. Some of the known issues with today's leading AI models include:

Explainability This is the ability to explain the reasoning and describe the data underlying an AI system's conclusions. Explainability is useful for:

- Giving users confidence that an AI system works well
- Safeguarding against bias of various kinds
- Adhering to regulatory standards or policy requirements
- Helping developers understand why a system works a certain way, assess its vulnerabilities, or verify its outputs
- Meeting society's expectations about how individuals are afforded agency in a decision-making process[30]

Today's AI is for the most part incapable of explaining the basis on which it arrives at any particular conclusion. Explanations are not always relevant, but in certain cases, such as medical decision making, they may be critical.

Bias and fairness Because machine-learning models are trained on existing datasets, such models are likely to encode any biases present in such datasets. (Bias should be understood here as a property of the data that is commonly regarded as societally undesirable.) For example, if a facial recognition system is primarily trained on images of individuals from one ethnic group, its accuracy at identifying people from other ethnic groups may be reduced.[31] Use of such a system could well lead to disproportionate singling out of individuals in those other groups. To the extent that these datasets reflect history, they will also reflect the biases embedded in history, and a machine-learning model based on such datasets will also reflect such biases.

Vulnerability to spoofing For many AI models, it is possible to tweak data inputs to fool them into drawing false conclusions. For example, in figure 1.1, the addition of a small amount of noise to the "panda" image results in its being classified as a gibbon with very high confidence.

Deepfakes AI provides the capability for generating highly realistic but entirely inauthentic audio and video imagery. Such capability has obvious implications for evidence presented in a courtroom and political deception.

Privacy Many LLMs are trained on data found on the internet rather indiscriminately, and such data often include personal information of individuals. When incorporated into LLMs, such information could be more publicly disclosed.

Overtrust If AI systems become a common presence in society, their novelty will inevitably diminish for users. The level of trust in computer outputs often increases with familiarity. On the other hand, skepticism about answers received from a system is necessary if one is to challenge the correctness of these outputs. As trust in AI grows due to reduced skepticism, there's a higher risk of overlooking errors, mishaps, and unforeseen incidents. One experiment recently showed that those with access to an AI-based coding assistant wrote code significantly less secure than those without an AI-based assistant—even though the former were more likely to believe they had written secure code.[32]

Hallucinations AI hallucinations refer to situations where an AI model generates results or answers that are plausible but nevertheless do not correspond to reality. That is, AI models can simply make things up seemingly from whole cloth. The results are plausible because they are constructed based on statistical patterns that the model has learned to recognize from its training data. But they may not correspond to reality because the model does not have an understanding of the real world.

FIGURE 1.1 A panda turns into a gibbon

"panda"
57.7% confidence

"gibbon"
99.3% confidence

Source: Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy, "Explaining and Harnessing Adversarial Examples," paper presented at International Conference on Learning Representations, San Diego, May 2015.

Out-of-distribution (OOD) inputs All machine-learning systems must be trained on a large volume of data. If the inputs to a system are substantially different from the training data—or out of distribution—the system may well draw conclusions that are more unreliable than if the inputs were like the training data.

Copyright violations Some AI-based models have been trained on large volumes of data that have been found online. These data have generally been used without the consent or permission of their owners, thereby raising important questions about appropriately compensating and acknowledging these data owners.

AI researchers are cognizant of issues such as those described above, and in many cases, work has been or is being done to develop corrective measures for them. Nevertheless, it's fair to say that in most cases, such defenses do not generalize very far beyond the specific problems against which these defenses are deployed.

Challenges of Innovation and Implementation

The primary challenge of bringing AI innovation into operation is risk management. It is often said that AI, especially machine learning, brings a new conceptual paradigm for how systems can exploit information to gain advantage, relying on pattern recognition in the broadest sense rather than on explicit understanding of situations that are likely to occur. Because it is new, the people who would make the decision to deploy AI-based systems do not have a good understanding of the risks that might accompany such deployment.

Consider, for example, artificial intelligence as an important approach for improving the effectiveness of military operations. Despite broad agreement by the military services and the US Department of Defense (DOD) that AI would be of great benefit, the actual integration of AI-enabled capabilities into military forces proceeds at a slow pace. One important reason for this outcome is that the DOD acquisition system has largely been designed to minimize the likelihood of programmatic failure, fraud, unfairness, waste, and abuse—in short, to minimize risk. In this environment, it is not surprising that the incentives at every level of the bureaucracy are aligned in that manner. For new approaches (like AI) to take root, a greater degree of risk acceptance may well be necessary.

Policy, Legal, and Regulatory Issues

THE FUTURE OF WORK

Large language models such as GPT-4 have already demonstrated how they can be used in a variety of diverse fields, including law, customer support, coding, and journalism. These demonstrations have led to concerns that the impact of AI on employment will be substantial, especially on jobs that involve knowledge work, but uncertainty abounds. What and how many present-day jobs will disappear or be created? Which tasks could best be handled by AI?

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 feel painful shifts in the short term.[33]

- AI is helping some workers to increase productivity and job satisfaction.[34] At the same time, other workers are already losing their jobs as AI—despite potentially underperforming humans—demonstrates adequate competence for business operations.[35]

- 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 jobs resulting from widespread AI deployment are not clear at this point, although historically the introduction of new technologies has not resulted in a long-term net loss of jobs.[36]

Research on foundational AI technologies is difficult if not impossible to regulate. . . . Regulation of specific applications of AI may be more easily implemented.

REGULATION OF AI

Governments around the world have been increasingly focused on establishing regulations and guidelines on AI. Research on foundational AI technologies is difficult if not impossible to regulate, especially when other nations have strong incentives to carry on regardless of actions taken by US policymakers. The same applies to voluntary restrictions on research by companies concerned about the competition. Regulation of specific applications of AI may be more easily implemented, in part because of existing regulatory frameworks in application domains such as health care, finance, and law.

The European Union is advancing comprehensive legislation that would provide for harmonized rules for the governance of artificial intelligence to mitigate new risks or negative consequences for individuals or society.[37] In the United States, more nascent federal and state discussions have advanced ideas such as the creation of a centralized AI agency, licensing of AI companies, and registration and transparency requirements of AI models.[38] Whether regulatory efforts will in the end prove consistent with technical realities remains to be seen.

NATIONAL SECURITY

AI is expected to have a profound impact on the military.[39] Weapons systems, command and control, logistics, acquisition, and training will leverage multiple AI technologies to operate more effectively and efficiently, at lower cost, and with less risk to friendly forces. The DOD is therefore dedicating billions of dollars to institutional reforms and research advancements aimed at integrating AI into its warfighting and war preparation strategies. Military officials recognize 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 heavily investing in AI capabilities.

The DOD is also cognizant of its obligation to proceed ethically with the development of AI capabilities; it has adopted a set of guiding principles that address responsibility, equity, traceability, reliability, and governability in and for AI.[40] An additional important concern, subsumed under these principles but worth calling out explicitly, is determining where the use of AI may or may not be appropriate, such as AI in nuclear command and control.

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 AI students joining the industry, particularly start-ups, is increasing at the expense of those pursuing academic careers and contributing to foundational AI research.

Many factors are contributing to this trend. One is that industry careers come with compensation packages that far outstrip those offered by academia. Academic researchers must obtain funding to pay for research equipment and personnel like staff scientists, technicians, and programmers. They must search for government grants, which are 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.

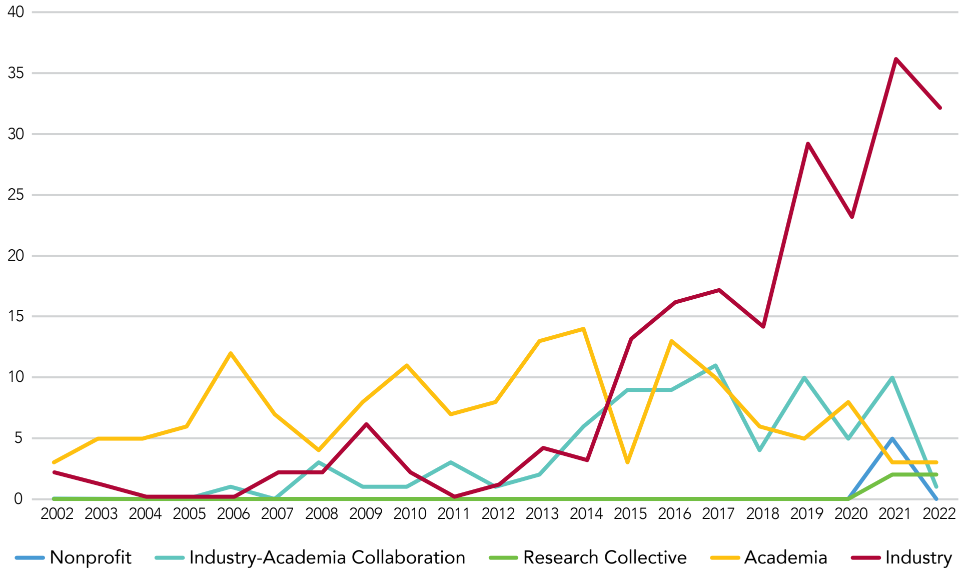
Industry often makes decisions more rapidly than government grant makers and imposes fewer regulations on the conduct of research. Large companies have the advantage of having research-supporting infrastructure in place, such as compute facilities and data storehouses.

One important consequence is that academic access to research infrastructure is limited, so US-based students are unable to train on state-of-the-art systems, at least not if their universities do not have access to the facilities of industry. Figure 1.2 shows how most significant machine-learning systems are now released by industry, with very few released by academic institutions, research collectives, or nonprofits.

At the same time, China's efforts to recruit top scientific talent offer further temptations for scientists to leave the United States. Although these efforts are often targeted toward ethnic Chinese in the United States—ranging from the well established to those finishing graduate degrees—China offers recruitment packages that promise benefits comparable to those available from private industry, such as high salaries, lavish research funding, and apparent freedom from bureaucracy.

All these factors are leading to an AI brain drain that does not favor the US research enterprise.

FIGURE 1.2 Number of significant machine-learning systems by sector



Source: Nestor Maslej et al., *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

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. Grand View Research, “Artificial Intelligence Market Size, Share, Growth Report 2023–2030,” accessed August 15, 2023, https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market.

3. Daniel Zhang et al., *The AI Index 2022 Report*, AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, March 2022, https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf.

4. Gené Teare, “Special Series Launch: The Promises and Perils of a Decade of AI Funding,” Crunchbase News, November 22, 2022, https://news.crunchbase.com/ai-robotics/ai-venture-cybersecurity-health-care-series-launch.

5. Shubham Sharma, “Report: AI Startup Funding Hits Record High of $17.9B in Q3,” VentureBeat, November 11, 2021, https://venturebeat.com/business/report-ai-startup-funding-hits-record-high-of-17-9b-in-q3.

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

7. 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.

8. 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.

9. Darian Woods and Adrian Ma, “The Semiconductor Founding Father,” December 21, 2021, in *The Indicator from Planet Money*, podcast produced by NPR, MP3 audio, 10:14, https://www.npr.org/transcripts/1066548023.

10. Ludvigsen, “The Carbon Footprint.”

11. Ludvigsen, “ChatGPT’s Electricity Consumption,” Medium, July 12, 2023, https://towardsdatascience.com/chatgpts-electricity-consumption-7873483feac4.

12. 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.

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

14. Gary Liu et al., “Deep Learning-Guided Discovery of an Antibiotic Targeting *Acinetobacter baumannii*,” *Nature Chemical Biology* (2023), https://doi.org/10.1038/s41589-023-01349-8.

15. Khari Johnson, “Hospital Robots Are Helping Combat a Wave of Nurse Burnout,” *Wired*, April 19, 2022, https://www.wired.com/story/moxi-hospital-robot-nurse-burnout-health-care.

16. Fish Site, “Innovasea Launches AI-powered Biomass Camera for Salmon,” August 17, 2023, https://thefishsite.com/articles/innovasea-launches-ai-powered-biomass-camera-for-salmon.

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

18. GateHouse Maritime, “Vessel Tracking Giving Full Journey Visibility,” accessed August 15, 2023, https://gatehousemaritime.com/data-services/vessel-tracking.

19. JD Supra, “Artificial Intelligence in Law: How AI Can Reshape the Legal Industry,” September 12, 2023, https://www.jdsupra.com/legalnews/artificial-intelligence-in-law-how-ai-8475732.

20. 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.

21. Rishi Bommasani et al., “On the Opportunities and Risks of Foundation Models,” arXiv, Cornell University, August 16, 2021, https://arxiv.org/abs/2108.07258.

22. Nestor Maslej et al., *The AI Index 2023 Annual Report*, AI Index Steering Committee, Stanford Institute for Human-Centered AI, April 2023, 81–98, https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

23. John Jumper et al., “Computational Predictions of Protein Structures Associated with COVID-19,” DeepMind, Google, August 4, 2020, https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19.

24. Amir Shanehsazzadeh et al., “Unlocking de Novo Antibody Design with Generative Artificial Intelligence,” bioRxiv, March 29, 2023, https://doi.org/10.1101/2023.01.08.523187.

25. Jonas Degrave et al., “Magnetic Control of Tokamak Plasmas through Deep Reinforcement Learning,” *Nature* 602 (2022): 414–19, https://doi.org/10.1038/s41586-021-04301-9.

26. Sébastien Bubeck et al., “Sparks of Artificial General Intelligence: Early Experiments with GPT-4,” arXiv, Cornell University, April 13, 2023, https://doi.org/10.48550/arXiv.2303.12712.

27. “Stop Talking about Tomorrow’s AI Doomsday When AI Poses Risks Today,” editorial, *Nature* 618 (June 2023): 885–86, https://www.nature.com/articles/d41586-023-02094-7.

28. Dina Bass, “Microsoft Invests $10 Billion in ChatGPT Maker OpenAI,” Bloomberg, January 23, 2023, https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai; Office of Management and Budget, Executive Office of the President, “Proposed Funding for the National AI Initiative as a Whole Is Measured in a Few Billion,” *Budget of the U.S. Government, Fiscal Year 2024*, 2023, 132–33, https://www.whitehouse.gov/wp-content/uploads/2023/03/budget_fy2024.pdf.

29. 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.

30. Royal Society, “Explainable AI: The Basics,” November 28, 2019, https://royalsociety.org/topics-policy/projects/explainable-ai.

31. Joy Buolamwini and Timnit Gebru, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification,” Conference on Fairness, Accountability, and Transparency,

Proceedings of Machine Learning Research 81 (February 2018): 1–15, https://www.media.mit.edu/publications/gender-shades-intersectional-accuracy-disparities-in-commercial-gender-classification

32. Neil Perry et al., “Do Users Write More Insecure Code with AI Assistants?,” arXiv, Cornell University, December 16, 2022, https://doi.org/10.48550/arXiv.2211.03622.

33. 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.

34. 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.

35. 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.

36. David Autor et al., “New Frontiers: The Origins and Content of New Work, 1940–2018,” National Bureau of Economic Research, August 2022, https://doi.org/10.3386/w30389.

37. European Commission, Proposal for a Regulation of the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, European Union, June 1, 2021, https://eur-lex.europa.eu/legal-content/EN/TXT.

38. Dylan Matthews, “The AI Rules That US Policymakers Are Considering, Explained,” Vox, August 1, 2023, https://www.vox.com/future-perfect/23775650/ai-regulation-openai-gpt-anthropic-midjourney-stable.

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

40. C. Todd Lopez, “DOD Adopts 5 Principles of Artificial Intelligence Ethics,” DOD News, US Department of Defense, February 25, 2020, https://www.defense.gov/News/News-Stories/article/article/2094085/dod-adopts-5-principles-of-artificial-intelligence-ethics.

02

BIOTECHNOLOGY AND SYNTHETIC BIOLOGY

KEY TAKEAWAYS

- Biotechnology is burgeoning, contributing around 5 percent to US GDP with a historical doubling time of about seven years.
- Synthetic biology is third-generation biotechnology, complementing domestication and breeding (the first generation) and gene editing (the second generation).
- The United States is struggling to grasp the scale of the bio-opportunity, the strategic ramifications unique to network-enabled biotechnologies, and the possibilities and perils of distributed biomanufacturing.

Overview

Biotechnology depends on molecular and cellular methods to realize breakthrough products or services.[1] As representative examples, the Stanford biotechnology community in 2023 pioneered the bioengineering of skin microbes to combat skin cancer,[2] realized industrial-scale translation for yeast-based brewing of essential medicines,[3] and achieved full resolution imaging of precursor synthetic cells, setting the stage for a “life race” akin to last century’s space race.[4]

Biotechnology-based products and services are already deployed at scale, having impacts equaling or exceeding those of more mature technologies due to the intrinsic power of biology. Yet leaders of one Fortune 100 company noted that biotechnology today is like a “snowflake on the tip of an iceberg.”[5] Stated differently, most of biotechnology hasn’t been imagined yet. This dual reality (i.e., *applications* enabled via immature and still-emerging *methods*) creates the potential for confusion or bad decisions.

The ancient Greek word *sunthesis* means "composition" or "a putting together." Synthetic biology thus focuses on fundamental methods that improve the composition and putting together of living systems, primarily at the molecular to cellular scales but increasingly at the tissue and microbial population levels. Building on genetic engineering, synthetic biology is not limited to genes as found in nature but whatever can be engineered and composed for specific purposes (e.g., an enzyme evolved by humans to catalyze carbon-silicon bonds).[6] Just as airplanes and rockets enabled humans to overcome some constraints of land and gravity, synthetic biology enables humans to develop living organisms beyond the constraints of lineage, such as petunias that emit light (i.e., nightlights that need watering instead of an electrical outlet).[7]

A 2020 National Academies report valued the US bioeconomy at about 5 percent of GDP, or more than $950 billion.[8] A 2020 McKinsey report identified four hundred synthetic biology projects currently in the R&D pipeline, estimating such innovations could add $4 trillion in direct economic impacts over the next ten to twenty years.[9] This projected pace of bioeconomic doubling over the next seven or so years tracks the historical record.[10] Biotech venture capital funding was $29.7 billion in 2022, the second-highest year on record, following the record $38.7 billion invested in 2021.[11]

Estimates of niche and still nascent synthetic biology markets vary widely, from $37 billion by 2028 to $100 billion by 2030.[12] A conservative estimate by the Congressional Research Service reported that US government research funding for synthetic biology increased from about $29 million in fiscal year 2008 to nearly $161 million in fiscal year 2022. Many first-generation synthetic biology companies have struggled or worse, suggesting that some ideas need revisiting or more support and that translation efforts would benefit from smarter strategies.[13]

In principle, anything that can be encoded in DNA could be grown when and where needed. In other words, biology can be regarded as the ultimate distributed manufacturing. For this reason, some have called for biology to be recognized as a *general-purpose technology* akin to computing, triggering associated calls for strategy and leadership.[14] As one representative far-reaching vision, in 2018, the Semiconductor Research Corporation outlined an ambitious twenty-year synthetic biology road map with its ultimate goal being to enable bottom-up construction of microprocessors.[15]

A Synthetic Biology Primer

DNA is both physical hereditary material and a digital code of life. DNA can be represented abstractly as four bases (A, C, T, and G). Particular sequences (i.e., orderings) of bases encode different living functions including biomolecules. Cells consist of encoded molecules and realize different behaviors and functions by producing the various molecules at the appropriate time and place.

DNA sequencing and synthesis are two fundamental technologies underlying synthetic biology. Sequencers are machines that read the precise series of As, Cs, Ts, and Gs encoding genetic information, while synthesizers write user-specified sequences of As, Cs, Ts, and Gs. The cost of sequencing a human genome has fallen from $10,000 to $600 in the last decade,[16] while the cost of gene synthesis has dropped a hundredfold from 2005 to 2015.

Improvements in DNA sequencing methods were jump-started and driven initially by government support for the initial human genome sequencing project. This public investment was sufficient to kick-start significant downstream market opportunities that

Most of biotechnology hasn't been imagined yet.

have since driven ongoing innovation and improvement. There has been no equivalent public support for improving DNA synthesis. Improvements in DNA synthesis have been sporadic and dependent on private capital, and the cost of commercially available gene-length DNA synthesis services has not improved significantly in the last six years.[17]

SARS-CoV-2's arrival in Switzerland in February 2020 illustrates the powerful potential of DNA sequencing and synthesis together, combined with the information transmission capabilities of the internet. Before the pandemic could naturally arrive, a researcher in China sequenced the virus's genome, uploaded a digital file representing that genome to the internet, from which a Swiss researcher downloaded the information, ordered the DNA, reconstructed the genome, and infected cells in the laboratory, accomplishing all this twelve days before the actual pandemic arrived over the Italian border.[18]

This example suggests that the "superpower" of the internet—the ability to rapidly move information—might usefully and ultimately recombine with the superpower of biology, namely the ability to grow and assemble complex objects locally. Stated differently, a DNA synthesizer is a "1-D printer," but the polymer it prints (i.e., DNA) in turn programs biomolecules that construct and assemble 3-D objects with atomic precision. DNA sequencers and synthesizers connected to the internet could thus routinely allow researchers to move viruses around the world faster than a pandemic can spread. Developed wisely, such capabilities could lead to biodefense and public health systems operating at light speed. Ignored or mismanaged, such capabilities could result in widespread access to bioterror capacities or worse.

Along with DNA reading and writing, synthetic biology is slowly advancing our ability to coordinate the composition of living systems. One line of work seeks to enable coordination of labor via reliable reuse of materials, measurements, and models. Example projects include developing standards for quantifying gene expression levels inside cells,[19] akin to how telegraph engineers long ago struggled to make and maintain communication systems using copper wire prior to the invention of the Ohm as a common unit of resistance.[20] Such fundamental research enabling coordination of labor allows many more to learn about, safely participate as citizens of, and benefit from the world of emerging biotechnologies and enables experts to realize products of increasing complexity more quickly. Such foundational research is almost entirely unsupported domestically at present but is increasingly a topic of discussion for international standards-setting bodies.

More ambitious projects seek to learn to construct artificial (or synthetic) cells entirely from scratch.[21] While there are many applications of such cells, the fundamental motivation is to make routine the engineering of living systems. For context, no natural cell used in any biotechnology process is fully understood. All natural cells require a significant number of essential genes whose encoded functions remain entirely unknown. Thus, contemporary biotechnology workflows remain Edisonian (i.e., tinker and test). By learning to construct cells from scratch, synthetic biologists are seeking to architect an operating system for life at its most fundamental level.

Key Developments

Synthetic Biology Applications

Being able to engineer and thus to modify existing cellular functions, synthetic biology has been able to make contributions to:

- **Medicine** DNA and RNA synthesis underlie all mRNA vaccines, including those for COVID-19. Synthetic biology also enables more sophisticated engineering of cell-based diagnostics and therapies, from bioengineered immune cells to microbiomes.[22]

- **Agriculture** Synthetic biology has been used to cultivate drought-resilient agricultural crops, enhance food security with indoor farming, offer plant- or cell-based meat cultivation, and improve food safety through easier tracing of contaminated products.

- **Manufacturing** Synthetic biology enables cells to be programmed as efficient, sustainable factories for medicinal drugs, fuels, and other useful substances.[23] One estimate expects "as much as 60 percent of the physical inputs to the global economy could, in principle, be produced biologically."[24]

- **Sustainability** Synthetic biology enables carbon-neutral and carbon-negative manufacturing. Developments in electrobiosynthesis (i.e., growing biomass from CO_2 and electricity)[25] and mycological manufacturing[26] are particularly compelling. Direct and indirect impacts of synthetic biology on biodiversity and conservation biology are gaining increasing attention.[27]

Artificial Intelligence in Synthetic Biology

In recent years, we have also seen computational methods enabled by artificial intelligence (AI) realize significant advances in predicting the three-dimensional shapes of proteins (one important class of biomolecules) from DNA sequence information. The specific shapes of proteins determine their function in the body.

Traditionally, determining protein shapes required the use of expensive experimental methods such as X-ray crystallography. However, in 2022, researchers used AI to predict the structures of more than two hundred million proteins from some one million species directly from their DNA sequence information, representing nearly every known protein on the planet.[28] About 35 percent of the predictions have been found to be as good as experimentally determined structures and perhaps an additional 45 percent are interesting enough to guide research. Similar methods are being developed to predict RNA structures.[29] Note, however, that structure alone may be insufficient to predict the function (i.e., the actual physical or biochemical behavior of the resulting biomolecule).[30] The success of such computational methods often depends on large sets of training data obtained (thus far) via decades of experimentation by far-flung research communities.

AI-based approaches are also being developed to aid in the design of genetic constructs.[31] Imagine a ChatGPT-like capacity enabling natural language requests that result in DNA-sequence designs for functional biomolecular systems. For example, "Hey Siri, get me a plasmid that will make *E. coli* smell like bananas when growing and wintergreen when dormant" could soon result in a well-designed DNA construct for synthesis.

Physics in Synthetic Biology

Computational methods in biology also depend on the validity of our mathematical representations of the physics of living systems. Such representations are well established for structural biology (i.e., how atoms are organized to comprise proteins and nucleic acids) but are less mature for cellular-scale systems (i.e., how molecules are organized to compose cells).[32] The application of colloidal hydrodynamic modeling to cellular-scale systems has recently enabled rational design of cellular-scale systems.[33]

Over the Horizon

Impact of Synthetic Biology

Future applications of synthetic biology may include:[34]

- Biomanufacturing of chemicals, solvents, detergents, reagents, plastics, electronic films, fabrics, polymers, agricultural products such as feedstock, crop protection solutions, food additives, fragrances, and flavors[35]
- Synthetic fuels that are energy dense enough for transportation produced by recycling carbon from sources such as cellulosic feed stocks, crops that make oil, and agriculture and municipal wastes[36]
- Nutrient-dense, drought-resistant crops that improve food and water security[37]
- Concrete that fixes carbon while curing and construction materials embedded with biomolecular functions that "heal" cracks[38]
- Biologically active paint that prevents biofouling of ship hulls or reduces pipeline corrosion[39]
- Biomining of critical minerals and bioengineered materials produced locally, contributing to more efficient, robust, and secure supply chains[40]
- A global biosecurity infrastructure that rapidly detects the emergence of pathogens anywhere on Earth and enables the rapid manufacturing of tailored vaccines, testing equipment, rapid therapeutics, and other treatments at the source of the outbreak within days[41]

Challenges of Innovation and Implementation

STRATEGIC VISION

Most discussions of and investment in biotechnology are understandably motivated by and focused on applications of biotechnology. However, from a governance and strategy perspective, advances in underlying methods that change what biotechnologies are possible are where strategies take root, where leverage begins, and where a strategic vision is most needed. US federal policy seeks to advance synthetic biology and biotechnology more broadly. Several building blocks are in place, including the following:

- The congressionally mandated National Engineering Biology Research and Development Initiative was established by Public Law 117-167, Section 10402—commonly known as the CHIPS Act of 2022.
- The National Biotechnology and Biomanufacturing Initiative was launched under Executive Order 14081 of September 2022.[42] It seeks whole-of-government action to increase domestic biomanufacturing capability, expand market opportunities for bio-based products, drive R&D, streamline regulation, and improve biosafety and biosecurity.[43]
- National Security Memorandum 15, "Countering Biological Threats, Enhancing Pandemic Preparedness, and Achieving Global Health Security" (NSM 15), signed in October 2022, establishes as a goal that the United States must "fundamentally transform its capabilities to protect our Nation from biological threats and advance pandemic preparedness and health security more broadly for the world."
- A report from the President's Council of Advisors on Science and Technology (PCAST), *Biomanufacturing to Advance the Bioeconomy*, was submitted to the president in December 2022 and is discussed below.[44]

Other proposals have gone further than the vision laid out in current federal policy. For example, the Special Competitive Studies Project calls for biotechnology moon shots to advance the underlying science and technology behind construction of fully synthetic cells; the alignment of incentives for biotechnology commercialization such as the local biomanufacture of medicines; and the building-out of infrastructure to support the biotechnology enterprise, including research and manufacturing facilities, data management policies, a skilled workforce, and international cooperation with likeminded nations.[45]

The National Academies of Sciences, Engineering, and Medicine reported that the United States has pioneered advancements in biotechnology. But their report, along with reports from others, also noted the emergence of significant competitors, China and Europe in particular, and therefore that any US lead cannot be taken for granted.[46]

The PCAST report mentioned above identified three challenges that must be addressed to ensure the United States maintains its leadership and fully exploits the benefits of the bioeconomy:

- The lack of an adequate US biomanufacturing capacity and workforce
- An outdated US regulatory process for many new bioproducts that can delay or stop their commercialization
- An integrated and overarching bioeconomy strategy to help guide federal agencies in managing the development and transfer of these biotechnologies toward social and economic advancements

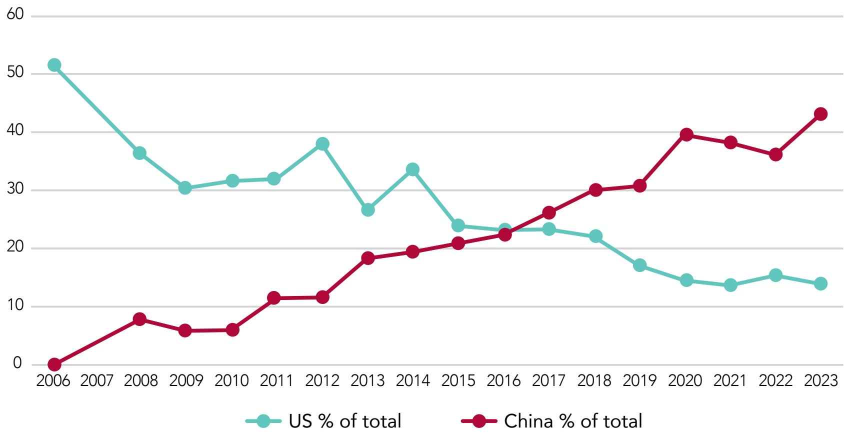
As an illustration of the first bullet above, consider the International Genetically Engineered Machines (iGEM) Competition, which involves teams from all over the world composed of self-funded students enrolled in high schools or institutions of higher learning, or from people working in community labs. The competition, which began at the Massachusetts Institute of Technology in 2003, gives each team a kit with a variety of genetic parts and asks them to use their own laboratories to create bioengineered organisms that address a local need or problem. About one hundred thousand students have participated in iGEM and many iGEM alumni are now leaders (e.g., the chair of the US Congress's National Security Commission on Emerging Biotechnology is an iGEM alum). In 2003, the competition had only American teams. But over the last decade, teams from China have outnumbered US teams by a factor of three, and in 2023 there were 175 teams from China compared to 56 from the United States (see figure 2.1). European teams have also been increasing their numbers.

STANDARDS SUPPORTING R&D AND TRANSLATION

One key early institution for synthetic biology was the Registry of Standard Biological Parts operated and funded by the iGEM Foundation, offering a collection of genetic parts used in the synthesis and assembly of new biological components, systems, and devices. The "standard" adjective in "Standard Biological Parts" means that any given part is compatible (to a limited degree) with other similarly standardized parts and can therefore be integrated into larger and more complex assemblies while still maintaining the compatibility format of the standard.[47] Users of the registry are encouraged, but not required, to contribute data and develop new parts to enhance the resource. Today, the Registry contains over twenty thousand parts.

This one physical assembly standard, however, is not enough. For example, biological data obtained in one laboratory needs to be usable across the entire synthetic biology community. Computational models of biological processes and organisms should be usable across the entire community to validate results. Such interoperability requirements

FIGURE 2.1 Percentage of iGEM teams from the US and China



Source: iGEM, Team List of iGEM Championship (data for 2008 through 2023); iGEM, Schools Partipating in iGEM 2006 (data for 2006).

often drive the need for standards that specify what data elements must be retained, what annotations need to be provided, and in what format they must be retained. These standards are necessary to ensure that biological data and computational models are usable across the entire community.

The effort to develop standards must include academia and actors from the private sector. The lines between research and specific applications are particularly fuzzy for synthetic biology in that private-sector firms support a substantial amount of research in the field. These firms include companies dedicated to exploring the commercial potential in synthetic biology, from start-ups to major pharmaceuticals, as well as firms that have historically been focused on information technology.

The US government could play a critical role in supporting technical standards underlying biotechnology. Traditionally, the US approach to standards development is more market driven than, for example, the European approach.[48] But given that the market is unlikely to develop standards that can support collaborative work in both academia and the private sector, a degree of government involvement in this enterprise would not be inconsistent with the intent expressed in the CHIPS Act of 2022 (discussed in chapter 8 on semiconductors) to support strategically important fields.

SUSTAINED R&D FUNDING

Although mRNA vaccines came into widespread public view in 2021, their history began some thirty years ago, with academia and industry both playing key roles in advancing the science.[49] These scientists improved the understanding of the mRNA pharmacology and made novel insights in immunology, laying the foundation for the next-generation mRNA vaccines. This history strongly suggests the need for "patient capital"—that is, investment in R&D that is sustained in times of ebb and flow in the pace of scientific advancement.

Also, from a competitiveness perspective, as the scope of bio-opportunities grows, the total amount of funding for basic research and translation must also grow. All of about forty Stanford faculty interviewed for this chapter clearly highlighted the issue of limited funding for foundational biotechnology research; one recent Nobel laureate reported that over 90 percent of her research projects remain unfunded.

Policy, Legal, and Regulatory Concerns

ENVIRONMENTAL AND SAFETY RISKS

New organisms not found in nature raise concerns about how they will interact with natural and human environments. For example, engineered cells in the human body can lead to unanticipated adverse effects. Bioengineered organisms that escape into the environment and possibly disrupt local food webs or displace natural species have long been a concern. Importantly, synthetic biology itself offers the possibility of bioengineering organisms from scratch that are incapable of escaping or evolving.[50] Such examples highlight how political and cultural concerns need not wait to be expressed and addressed; governments could facilitate and underwrite more active realization of the public interests in emerging biotechnologies.

NATIONAL SECURITY, PANDEMIC PREPAREDNESS, AND PUBLIC SAFETY CONSIDERATIONS

As the science and technology of synthetic biology becomes increasingly available to state and non-state actors, there are legitimate concerns that malicious actors will create organisms harmful to people or the environment. For example, polio, horsepox, SARS-CoV-2, and influenza have been synthesized from scratch in laboratories.

The US government does have an explicit policy for the oversight of research in the biological sciences, known as "dual use research of concern,"[51] focused on certain high-consequence pathogens and toxins. The policy is intended to preserve the benefits of such research while minimizing the risk of misuse of the knowledge, information, products, or technologies associated with it. Nevertheless, the policy covers only research funded or conducted by the US government, research involving one or more specified agents on a US government list, or one of several specific types of experiments. Moreover, despite growing concerns, such research is not per se illegal under international law (the Biological Weapons Convention) as long as it is consistent with the general-purpose criterion in Article I of the convention, leaving some to argue that education is the only way to significantly reduce the likelihood that such work will be conducted.[52]

GAPS IN REGULATORY REGIME

The PCAST report highlighted the inadequacies in the current regulatory process for approving biotech products. The current regulatory regime is the Coordinated Framework for the Regulation of Biotechnology, which splits biotech regulation among three different federal agencies, the Environmental Protection Agency (EPA), the US Department of Agriculture (USDA), and the Food and Drug Administration (FDA). The oversight is based on the end products' characteristics and unique features rather than on their production method.[53] However, some have voiced concern over whether the Coordinated Framework is sufficient given the increasingly complex, novel, and broad applications of synthetic biology that "go beyond contained industrial uses and traditional environmental release."[54]

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. Often classified as potential nonphysical impacts, the effects of synthetic biology when considering these religious concerns are difficult to predict in advance.

In the words of a Wilson Center report on this topic, concerns about nonphysical impacts are primarily "concerns about the appropriate attitude to adopt toward ourselves and the rest of the natural world."[55] The report notes that these 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." The report also notes that many people disagree about "whether a particular activity threatens these values, how we should reduce nonphysical harm, who should be responsible and what may be sacrificed along the way. . . . We do not always agree about what counts as a nonphysical harm, because we disagree about what is human well-being . . . [and this is because we embrace] different ethical frameworks."

In short, policymakers will have to be aware of and able to navigate issues and aspects of synthetic biology such as those described above if they are to help guide the development of the field and the increasing diversity of the resulting biotechnologies.

NOTES

1. International Union of Pure and Applied Chemistry, *Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"), comp. by A. D. McNaught and A. Wilkinson (Oxford: Blackwell Scientific Publications, 1997), online version (2019–) created by S. J. Chalk, https://doi.org/10.1351/goldbook.

2. 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.

3. SynBioBeta, "Antheia's Monumental Commercial Achievement," August 24, 2023, https://www.synbiobeta.com/read/antheias-monumental-commercial-achievement.

4. Anton Jackson-Smith, "Integration and Measurement for Building Synthetic Cells" (doctoral dissertation, Stanford University, 2023).

5. Anonymous personal communication to Professor Drew Endy, Stanford University, 2014. Drew Endy is a member of the SETR faculty council.

6. S. B. Jennifer Kan et al., "Directed Evolution of Cytochrome c for Carbon-Silicon Bond Formation: Bringing Silicon to Life," *Science* 354, no. 6315 (November 2016): 1048–51, https://doi.org/10.1126/science.aah6219.

7. Tatiana Mitiouchkina et al., "Plants with Genetically Encoded Autoluminescence," *Nature* 38 (April 2020): 944–46, https://doi.org/10.1038/s41587-020-0500-9; Animal and Plant Health Inspection Service, "Aphis Issues Regulatory Status Review Response: Light Bio Petunia," US Department of Agriculture, September 6, 2023, https://www.aphis.usda.gov/aphis/newsroom/stakeholder-info/sa_by_date/sa-2023/rsr-light-bio-petunia.

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

9. Michael Chui et al., "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. Planetary Technologies, "Bioeconomy Dashboard," last modified 2023, https://www.planetarytech.earth/bioeconomy-dashboard-1.

11. Brian Gormely, "Tales from the 2022 Biotech VC Fundraising Trail," *Wall Street Journal*, December 22, 2022, https://www.wsj.com/articles/tales-from-the-2022-biotech-vc-fundraising-trail-11671667063.

12. Congressional Research Service, *Synthetic/Engineering Biology: Issues for Congress*, September 30, 2022, 2, https://crsreports.congress.gov/product/pdf/R/R47265/2.

13. Reuters, "Biotech Firm Amyris Files for Bankruptcy in US," 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.

14. Abigail Kukura 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.

15. Semiconductor Research Corporation, "SemiSynBio Consortium and Roadmap Development," accessed September 18, 2023, https://www.src.org/program/grc/semisynbio/semisynbio-consortium-roadmap.

16. Emily Mullin, "The Era of Fast, Cheap Genome Sequencing Is Here," *Wired*, September 29, 2022, https://www.wired.com/story/the-era-of-fast-cheap-genome-sequencing-is-here.

17. Planetary Technologies, "Bioeconomy Dashboard," accessed September 18, 2023, https://www.planetarytech.earth/bioeconomy-dashboard-1.

18. Tran Thi Nhu Thao et al., "Rapid Reconstruction of SARS-CoV-2 Using a Synthetic Genomics Platform," *Nature* 582 (2020): 561–65, https://doi.org/10.1038/s41586-020-2294-p.

19. Jason R. Kelly et al., "Measuring the Activity of BioBrick Promoters Using an In Vivo Reference Standard," *Journal of Biological Engineering* 3, no. 4 (2009), https://doi.org/10.1186/1754-1611-3-4.

20. Graeme J. N. Gooday, *The Morals of Measurement: Accuracy, Irony, and Trust in Late Victorian Electrical Practice* (Cambridge:

Cambridge University Press, 2004), https://doi.org/10.1017/CB09780511550690.

21. Pauline van Nies et al., "Self-Replication of DNA by Its Encoded Proteins in Liposome-Based Synthetic Cells," *Nature Communications* 9, no. 1583 (2018), https://doi.org/10.1038/s41467-018-03926-1.

22. Mario Aguilera, "Researchers Engineer Bacteria That Can Detect Tumor DNA," University of California–San Diego, August 10, 2023, https://today.ucsd.edu/story/researchers-engineer-bacteria-that-can-detect-tumor-dna.

23. Amy Webb and Andrew Hessel, "Chapter Six. The Biological Age," in *The Genesis Machine: Our Quest to Rewrite Life in the Age of Synthetic Biology* (New York: Public Affairs, Hachette Book Group, 2022): 110–35.

24. Chui, "The Bio Revolution."

25. Robert F. Service, "Chemists Convert Electricity into the Fuel That Powers the Body's Cells," *Science*, August 22, 2023, https://www.science.org/content/article/chemists-convert-electricity-fuel-powers-body-s-cells.

26. PR Newswire, "MycoWorks to Open World's First Commercial-Scale Fine Mycelium™ Plant in September," August 7, 2023, https://www.prnewswire.com/news-releases/mycoworks-to-open-worlds-first-commercial-scale-fine-mycelium-plant-in-september-301894284.html.

27. Nicholas B. W. Macfarlane et al., "Direct and Indirect Impacts of Synthetic Biology on Biodiversity Conservation," *iScience* 25, no. 11 (October 2022): 105423, https://doi.org/10.1016/j.isci.2022.105423; "Superorganism," Superorganism, accessed September 18, 2023, https://www.superorganism.com.

28. 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.

29. Isabel Swafford, "Stanford Machine Learning Algorithm Predicts Biological Structures More Accurately Than Ever Before," Stanford News, August 26, 2021, https://news.stanford.edu/2021/08/26/ai-algorithm-solves-structural-biology-challenges.

30. Derek Lowe, "Making Up Proteins," *Science*, January 30, 2023, https://www.science.org/content/blog-post/making-proteins.

31. PR Newswire, "Ginkgo Bioworks and Google Cloud Partner to Build Next Generation AI Platform for Biological Engineering and Biosecurity," August 29, 2023, https://www.prnewswire.com/news-releases/ginkgo-bioworks-and-google-cloud-partner-to-build-next-generation-ai-platform-for-biological-engineering-and-biosecurity-301912283.html.

32. Akshay J. Maheshwari et al., "Colloidal Hydrodynamics of Biological Cells: A Frontier Spanning Two Fields," *Physical Review Fluids* 4, no. 11 (November 2019), https://doi.org/10.1103/PhysRevFluids.4.110506.

33. Akshay J. Maheshwari et al., "Engineering tRNA Abundances for Synthetic Cellular Systems," *Nature Communications* 14, no. 2594 (2023), https://doi.org/10.1038/s41467-023-40199-9.

34. Christopher A. Voigt, "Synthetic Biology 2020–2030: Six Commercially-Available Products That Are Changing Our World," *Nature Communications* 11, no. 6379 (December 2020), https://doi.org/10.1038/s41467-020-20122-2.

35. US Department of Defense, "DOD Approves $87 Million for Newest Bioindustrial Manufacturing Innovation Institute," October 20, 2020, https://www.defense.gov/News/Releases/Release/Article/2388087/dod-approves-87-million-for-newest-bioindustrial-manufacturing-innovation-insti.

36. Emily Aurand et al., *Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy* (Emeryville, CA: Engineering Biology Research Consortium, 2019), https://roadmap.ebrc.org/wp-content/uploads/2020/06/Roadmap-FINAL-18June2019-_Corrected.pdf.

37. Kukura, "National Action Plan," 1.

38. S. Udhaya et al., "Experimental Study on Bio-Concrete for Sustainable Construction," *Materials Today: Proceedings* (April 2023), https://doi.org/10.1016/j.matpr.2023.03.676.

39. Santosh Kumar et al., "Nanocoating Is a New Way for Biofouling Prevention," *Frontiers* 3, no. 771098 (November 2021), https://doi.org/10.3389.fnano.2021.771098.

40. Kukura, "National Action Plan," 1.

41. Kukura, "National Action Plan," 1.

42. White House, Executive Order on Advancing Biotechnology and Biomanufacturing Innovation for a Sustainable, Safe, and Secure American Bioeconomy, Presidential Actions, Briefing Room, White House, September 12, 2022, https://www.whitehouse.gov/briefing-room/presidential-actions/2022/09/12/executive-order-on-advancing-biotechnology-and-biomanufacturing-innovation-for-a-sustainable-safe-and-secure-american-bioeconomy.

43. Congressional Research Service, *The Bioeconomy: A Primer*, September 19, 2022, 12, https://crsreports.congress.gov/product/pdf/R/R46881.

44. President's Council of Advisors on Science and Technology, *Report to the President: Biomanufacturing to Advance the Bioeconomy*, December 2022, 2, https://www.whitehouse.gov/wp-content/uploads/2022/12/PCAST_Biomanufacturing-Report_Dec2022.pdf.

45. Kukura, "National Action Plan."

46. Congressional Research Service, *The Bioeconomy*; Executive Order on Advancing Biotechnology; Kukura, "National Action Plan," 18–19.

47. iGEM Foundation, "Help: Synthetic Biology," accessed August 15, 2022, http://parts.igem.org/Help:Synthetic_Biology.

48. Daniel S. Hamilton, "Promoting U.S.-EU Coordination and Cooperation on Technology Standards: Recommendations for Actions," Transatlantic Leadership Network, Trade and Technology Working Group, January 2021, https://www.transatlantic.org/wp-content/uploads/2022/03/TTC-tech-standards-January-2021.pdf.

49. Rein Verbeke et al., "Three Decades of Messenger RNA Vaccine Development," *Nano Today* 28, no. 100766 (October 2019), https://doi.org/10.1016/j.nantod.2019.100766.

50. Jonathan Calles et al., "Fail-Safe Genetic Codes Designed to Intrinsically Contain Engineered Organisms," *Nucleic Acids Research* 47, no. 19 (November 2019): 10439–51, https://doi.org/10.1093/nar/gkz745; Akos Nyerges et al., "A Swapped Genetic Code Prevents Viral Infections and Gene Transfer," *Nature* 615 (2023): 720–27, https://doi.org/ 10.1038/s41586-023-05824-z.

51. US Department of Health and Human Services, *United States Government Policy for Oversight of Life Sciences Dual Use*

Research of Concern, March 28, 2012, https://www.phe.gov/s3/dualuse/Documents/us-policy-durc-032812.pdf.

52. Tatyana Novossiolova et al., "Addressing Emerging Synthetic Biology Threats: The Role of Education and Outreach in Fostering Effective Bottom-Up Grassroots Governance," in *Emerging Threats of Synthetic Biology and Biotechnology*, ed. Benjamin D. Trump et al. (Dordrecht, NL: Springer Nature, 2021), 81–102, https://doi.org/10.1007/978-94-024-2086-9_6.

53. US Department of Agriculture, "About the Coordinated Framework," accessed August 15, 2023, https://usbiotechnologyregulation.mrp.usda.gov/biotechnologygov/about/about.

54. National Academies of Sciences, Engineering, and Medicine, "Conclusions and Recommendations," in *Preparing for Future Products of Biotechnology* (Washington, DC: National Academies Press, January 1, 2017) 171–86, https://www.ncbi.nlm.nih.gov/books/NBK442201.

55. 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.

03

CRYPTOGRAPHY

KEY TAKEAWAYS

- Cryptography is essential for protecting information but will never be enough to secure cyberspace.
- Cryptocurrencies have received a great deal of media attention, but they are not the most important issue in cryptography today.
- Cryptocurrencies use blockchain technology, but they are not the same; blockchain has many other important and promising applications.

Overview

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 mathematics to protect data from being altered or accessed inappropriately.[1] We are typically unaware that many of our day-to-day interactions with computers and the internet involve cryptography, from securing our online shopping to protecting our cell phone calls.

Cryptography is often invisible, but it is essential for most internet activities such as messaging, e-commerce, banking, or even simple internet browsing. Yet cryptography alone will never be enough to ensure the confidentiality, integrity, or availability of information. Inherent vulnerabilities in the software code that underpins all our internet-connected devices and the strong incentives for bad actors—from criminals to nation states—to engage in cyberattacks that exploit human and technical vulnerabilities help to explain why cybersecurity will be an ongoing challenge.

Cryptography Basics: Public Keys, Private Keys, and Hashes

Here's an example: Drew has a private message intended only for Taylor. To keep it confidential, she scrambles (encrypts) the message using an encryption algorithm and transmits the scrambled message to Taylor as ciphertext. When Taylor receives the ciphertext, he unscrambles (decrypts) it to reveal what it originally said. This piece of decrypted text is known as the plaintext. Along comes Ellen, a third-party eavesdropper who wants to see the plaintext, so she must use any means at her disposal to break the cryptographically provided protection.

An example of an encryption algorithm is the shift cipher. Each letter in the plaintext is replaced by a letter that is some fixed number N of positions later in the alphabet. For example, if $N = 2$, Drew substitutes an A in the plaintext with a C in ciphertext, B in plaintext with D in ciphertext, and so on. If $N = 3$, then Drew substitutes A in plaintext with D in ciphertext. To decrypt the ciphertext, Taylor must know that Drew is using the shift cipher and must also know the value of N so that he can invert it. For example, knowing that $N = 2$, he knows to write down A when he sees C in the ciphertext. (Note that modern encryption algorithms are more sophisticated and secure than what has been presented here; they are also harder to explain.)

In this scenario, both Drew and Taylor must share a secret piece of information—the cryptographic key, which is a string of numbers needed both to encrypt and to decrypt the message. Drew and Taylor must also know that the algorithm is the shift cipher. If Ellen somehow learns both of those facts, Ellen can decrypt the message as well. This type of encryption algorithm—of which the shift cipher is an example—is known as symmetric cryptography, or secret-key cryptography. It requires a secure key distribution, which is a method of distributing secret keys to all parties who should have them—but preventing those who shouldn't from obtaining them.

Symmetric key cryptography proved to be inconvenient and awkward because it requires in-person, physical effort ahead of the first secure communication to be had between the communicating parties, which makes it hard to talk to new people over the internet. In the 1970s, Stanford professor Martin Hellman and Whitfield Diffie codeveloped a technique known as asymmetric cryptography or public-key cryptography. Public-key cryptography relies on a public key for encrypting messages that is freely available to everyone, which means it can be widely distributed even over insecure channels. However, decrypting a message requires a private key that is held only by the authorized party (see figure 3.1).[2] Although it is theoretically possible to derive a private key from a public key, that process (if well designed) would take much too long for practical purposes (e.g., it would take longer than the age of the universe). It is this essential property that is placed at risk by quantum computing, as discussed below.

The mathematics of cryptography also underlie the creation of secure hashes. A hash is designed to accept a message of any length and compute a unique fixed-length string of numbers—called the hash value—corresponding to that message. Hashes have two key properties. First, it is extremely difficult to find another message that results in the same string of numbers. Second, if all you have is the string of numbers, it is infeasible to recover the original message.

Using a secure hash function, the sender can use public-key cryptography to provide assurances of integrity—information that cannot be tampered with or altered in any way—and identity, in that the originator of the message is who he or she claims to be.

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.[3] Alice then sends the message and its digital signature to Bob (the receiver).

Upon receipt of the message, Bob 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 also that Alice was the party who sent it, since only Alice could have used Alice's 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 Servers 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.

Blockchain

Blockchain is a cryptographic technology for creating distributed ledgers in the computing cloud. A blockchain records transactions so that they cannot be altered retroactively without detection. Because the entire blockchain can be distributed over thousands of computers, it is always accessible; anyone can deploy an application for it, and no one can prevent any such deployment. Moreover, anyone can interact with this application, and no one can prevent such an interaction. Finally, data cannot be erased. Later transactions may indicate that corrections are necessary, but the original data remain.

A blockchain can be visualized as a chain of blocks where each block contains a single transaction and a cryptographic hash of the previous block, creating a chain in which every block except the first is linked to the previous block. As more transactions occur, the blockchain gets longer because more blocks are added to the chain.

FIGURE 3.1 How public-key cryptography works

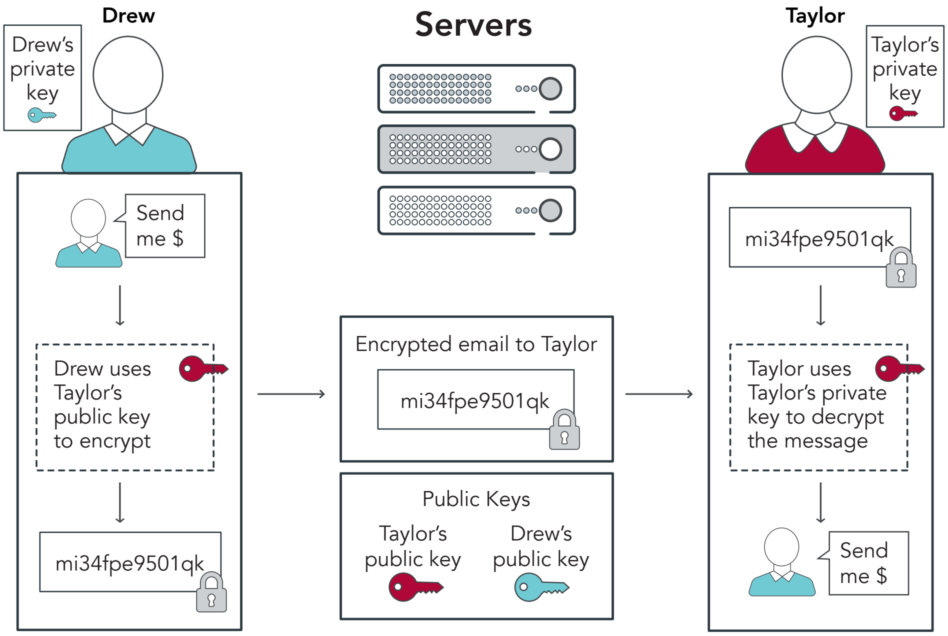
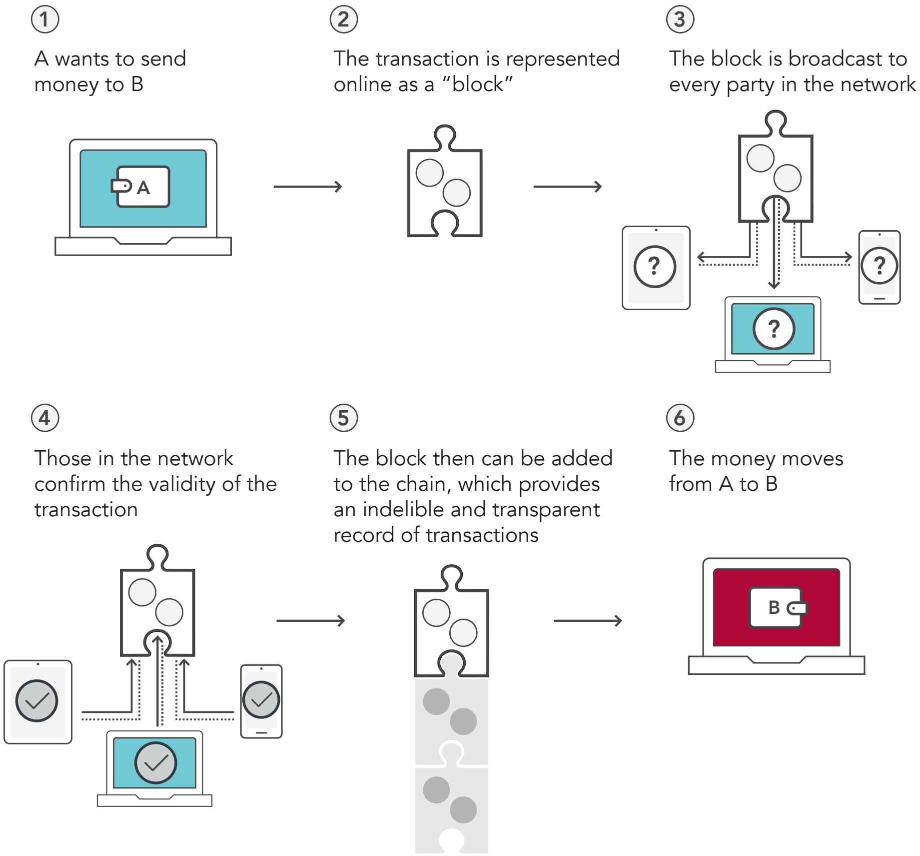


FIGURE 3.2 How a blockchain manages transactions



The distributed nature of blockchain also increases security. A new transaction is broadcast to every party in the network, each of which has a replica of the entire blockchain (see figure 3.2). Each party tries to validate the new transaction. It could happen that these replicas may not be fully synchronized; some might have received the new transaction while others did not. To ensure that all replicas are identical, blockchains have mechanisms for coming to consensus 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. These are computer programs that are always available and whose execution cannot be reversed—once a smart contract processes an incoming request, that processing cannot be reversed. Smart contracts can be used to implement financial instruments, to record ownership of digital assets, and to create marketplaces where people can buy and sell assets. Smart contracts are composable—one smart contract can use another—thus creating a vibrant ecosystem of innovation where one project can make use of a service developed by another project. Once deployed, they are available forever, running whenever someone interacts with them. By contrast, cloud computing applications are inherently transient—as soon as the application developer stops paying the cloud fees, the cloud provider kills the application.

Key Developments

A Host of Blockchain Applications

Blockchain technology was developed decades ago but has recently been used for a variety of applications. All those listed below have been implemented in some form and are operational today, though perhaps not on particularly large scales.

Time-stamping and data provenance Because data written to a blockchain cannot be modified or removed, blockchains provide a good mechanism for data provenance and time-stamping. An artist or an author who creates a new work of art can post a hash of the work to the chain, thereby proving the time at which the object was created. If later someone else claims authorship of the creation, the artist can point to the chain to prove its provenance.

Identity management A blockchain stores all the data from a person's important documents—diplomas, health-care and financial records, tax returns, birth certificate—in encrypted form. These original records are saved digitally, signed by their original providers, and, when made available through the blockchain, provided with provenance and time-stamping. Blockchain also facilitates selective revelation: upon request, the person can authorize release of data only to the minimal extent necessary to satisfy the request. For example, people can prove that their age is above some legal minimum, like twenty-one, but not have to reveal their date of birth. A person can allow a health-care researcher to look at her records for specific data—for example, whether she has ever had an abortion—without revealing her name. Applications of blockchain for identity management, such as SpruceID, are already being deployed.[4]

Supply chain management Blockchain can provide a transparent and secure way to track the movement of goods and their origin and quantity. This can be particularly valuable for high-value industries, such as the diamond industry; industries with significant counterfeit issues, such as luxury goods; or industries where the true source of goods is important, such as organic or vegan food. Blockchain can greatly simplify the job of forensic accountants trying to trace transactions.

Transactional records Many kinds of transactional records can be stored on a blockchain, thereby streamlining the process of buying and selling items by reducing fraud, increasing transparency, reducing paperwork, and making the process more efficient.

Cryptocurrencies Cryptocurrencies are digital instruments that many people use as a medium of exchange. Well-known cryptocurrencies include Bitcoin, Ethereum, Avalanche, and Polygon, each with 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, or real dollars, and generally use a cryptocurrency exchange to do so. Such exchanges are regulated financial institutions that transact in investments rather than currency.

Secure Computation

The field of cryptography has also expanded in scope to include secure computation, a well-established subfield that enables multiple parties to contribute inputs to a function that they jointly compute in such a way that the specific 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. Secure computation also allows users to prove they possess knowledge of a statement without having to disclose the actual content of that statement.

To illustrate secure computation, consider the problem of determining the collective wealth of three people while keeping the individual wealth of each person secret. Alice chooses a large random number and in secret adds her wealth to that number. Alice then gives the sum to Bob privately, who adds his wealth secretly to the number received from Alice. Bob secretly passes the total to Charlie, who does the same computation and then passes the result to Alice. Alice then in secret subtracts her original random number from the number received from Charlie and reveals the result to everyone else. That revealed number is the sum of each party's wealth but at no time does anyone learn of anyone else's wealth.[5]

This example is oversimplified (in fact, there is a subtle flaw in the procedure described). It's not exactly how a real-world secure computation works, but it suggests how computation on secret data might be accomplished. True secure computation protocols use more complex mathematics to defend against malicious behavior and to guarantee the privacy of each person's input during the computation process.

Applications of secure computation allow data analytics to be performed on aggregated data without disclosing the data associated with any individual element of the dataset. Banks can detect fraud without violating the privacy of individual customers. A group of workers can calculate their average salary without revealing their colleagues' personal pay. A Stanford system called Prio allows for a network of connected computers to work together to compute statistics, with clients holding their individual data privately.[6] This was deployed, for example, on mobile phones during COVID to calculate how many people were exposed to COVID in aggregate, without learning who was exposed.

Zero-Knowledge Proofs

A zero-knowledge proof is a cryptographic method that allows Paul (the prover) to prove to Vivian (the verifier) that Paul knows a specific piece of information without revealing to Vivian any details about that information. The term "zero-knowledge" indicates that Vivian gains zero new knowledge about the information in question, apart from the fact that what Paul is saying is true.

Consider a simplified example that demonstrates the logic: two people dealing with a locked safe. Let's say Paul wants to prove to Vivian that he knows the combination to the safe, but he doesn't want to reveal the combination to Vivian. With a zero-knowledge proof, Paul can convince Vivian that he knows the combination without exposing the combination itself.

To do so, Vivian writes something on a piece of paper and does not show it to Paul. Together, they put the paper into the safe and spin the combination lock. Vivian now challenges Paul to say what is on the paper. Paul responds by asking Vivian to turn around (so that Vivian cannot see Paul) and then enters the combination of the safe, opens it, looks at the paper and returns it to the safe, and closes it. When Vivian turns around, Paul tells her what was on the paper. Paul has thus shown Vivian that he knows the combination without revealing to Vivian anything about the combination.

In practice, of course, zero-knowledge proofs are more complex, yet they already have seen real-world implementations:

Banking A buyer may wish to prove to a seller the possession of sufficient funds for a transaction without revealing the exact amount of those funds. This capability has been implemented in the Zcash cryptocurrency.[7]

Provenance for digital images Cameras can provide a digital signature for every photo capturing an image and information about the time, date, and location. But such photos can then be digitally cropped, resized, or converted from color to black-and-white. Zero-knowledge proofs have been implemented in the standards of the Coalition for Content Provenance and Authenticity to ensure that the original photo was properly signed and that only permissible edits were made to the original without having to trust the editing software that was used.[8]

Cooperative tracking and verification of numbers of tactical nuclear warheads A zero-knowledge proof methodology has been developed to cooperatively provide updates on the movement and status changes of warheads in accordance with a political agreement to do so without revealing other sensitive information. This approach has not yet been implemented in any real arms control agreement, but its feasibility has been demonstrated in principle.[9]

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 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.

A second point is that widespread deployment will require confidence that proposed innovations will work as advertised. That is, 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. Expecting policymakers, consumers, and regulators to place their trust in these applications will be challenging.

Challenges of Innovation and Implementation

Although cryptography is fundamentally a mathematical discipline, it requires both human talent and substantial computing resources to examine the efficiency of new techniques, write software that is computationally expensive 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 that cryptographically enabled techniques and approaches could provide.

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 a factor of sixty). Thus, research faculty often tend to 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).

Policy, Legal, and Regulatory Issues

As a rule, public policy considerations are application specific; there has been no push to regulate basic research in cryptography for several decades.

Quantum-resistant algorithms are expected to be widely available by the time quantum computing comes online.

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.[10]

CRYPTOCURRENCY REGULATORY CONCERNS

Particularly considering the 2023 FTX trading scandal, in which the FTX cryptocurrency exchange went bankrupt and founder Sam Bankman-Fried was charged with fraud, many have questioned the extent to which cryptocurrencies should be exchangeable for national currency and whether they are better regulated as investment instruments or as currency. The lack of a regulatory framework for cryptocurrency affects many American users, consumers, and investors who are often confused about the basic workings of cryptocurrencies and their markets.

ENERGY CONSUMPTION

Bitcoin, an older and today the dominant cryptocurrency, consumes an enormous amount of energy; Bitcoin mining uses more energy than the Netherlands.[11] For this reason, newer blockchains—notably Ethereum—are designed to use far less energy, and today, Ethereum's annual energy use is less than 1/10,000 of YouTube's annual consumption. But Ethereum's market capitalization is less than half that of Bitcoin, and whether any less energy-intensive cryptocurrency will displace Bitcoin remains to be seen.

QUANTUM COMPUTING AND CRYPTOGRAPHY

Current public-key cryptography is based on the long times required with today's computers to derive a private key from its public-key counterpart. When realized, quantum computing (discussed more fully in chapter 8 on semiconductors) 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 United States 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. The first is that support for the transition to a quantum-resistant encryption environment should continue with urgency and focus.

A second issue is that messages protected by pre-quantum cryptography will be vulnerable in a post-quantum world. If those messages had been saved by adversaries (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.[12]

CENTRAL BANK DIGITAL CURRENCIES AND THE EROSION OF US FINANCIAL INFLUENCE

A central bank digital currency (CBDC) is a type of cryptography-based digital currency issued and regulated by a country's central bank, with legal tender status and value equivalent to the country's traditional currency—that is, digital assets backed by central banks. A CBDC can be designed with any number of the functional characteristics of cryptocurrencies and thus can be regarded as a "national cryptocurrency." However, a CBDC could be implemented in a centralized manner to improve performance and efficiency instead of using distributed blockchain technology.

An important benefit of a CBDC is the marriage of convenience and lower costs of digital transactions—by cutting out middlemen—and the regulatory oversight of traditional banking. In 2021, nearly six million Americans had no access to a bank account. Lower transaction costs would improve financial inclusion and enable many more people to have access to a well-regulated financial system. Those lower costs would also apply to cross-border transactions, therefore reducing the costs of international commerce.

The United States is considering issuing its own CBDC.[13] Although the dollar is the currency most used in cross-border transactions, the development of CBDCs by others could reduce global dependence on the dollar and on a financial infrastructure largely controlled today by the United States (e.g., SWIFT). This could significantly undermine the effectiveness of US economic sanctions and other financial tools. Today, more than ninety nations are researching, piloting, or deploying CBDCs, with several already testing cross-border transactions. China is the first major country to deploy a CBDC widely within its own economy, the digital yuan.[14]

NOTES

1. National Institute of Standards and Technology, "Cryptography," accessed August 15, 2023, https://www.nist.gov/cryptography.

2. Whitfield Diffie and Martin Hellman, "New Directions in Cryptography," *IEEE Transactions on Information Theory* IT-22, no. 6 (November 1976): 644–54.

3. In this context, encrypting the hash value simply means running the encryption algorithm using as the input key a string of numbers that just happen to be Alice's private key. 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.

4. https://www.spruceid.com.

5. This example is inspired by Keyless Technologies, "A Beginner's Guide to Secure Multiparty Computation," Medium, February 22, 2020, https://medium.com/@keylesstech/a-beginners-guide-to-secure-multiparty-computation-dc3fb9365458.

6. See "Prio," Stanford University, accessed September 25, 2023, https://crypto.stanford.edu/prio.

7. Zcash, "What Are Zero-Knowledge Proofs?," accessed August 30, 2023, https://z.cash/learn/what-are-zero-knowledge-proofs.

8. Trisha Datta and Dan Boneh, "Using ZK Proofs to Fight Disinformation," Medium, September 29, 2009, https://medium.com/@boneh/using-zk-proofs-to-fight-disinformation-17e7d57fe52f.

9. Miles A. Pomper et al., *OP55: Everything Counts: Building a Control Regime for Nonstrategic Nuclear Warheads in Europe*, CNS Occasional Paper Series, James Martin Center for Nonproliferation Studies, May 10, 2022, https://nonproliferation.org/op55-everything-counts-building-a-control-regime-for-nonstrategic-nuclear-warheads-in-europe.

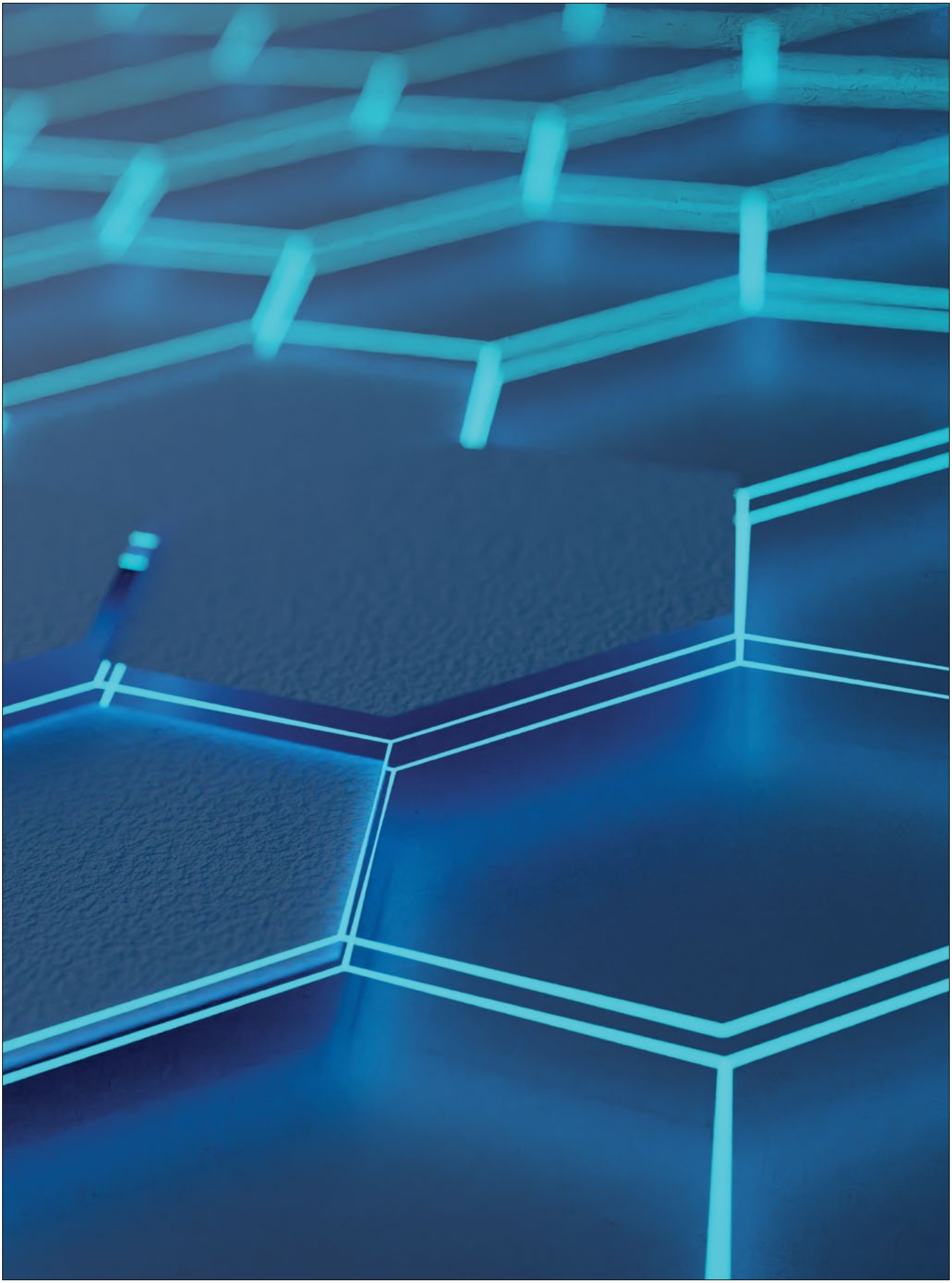
10. US Department of Justice, "Attorney General William P. Barr Delivers Keynote Address at the International Conference on Cybersecurity," July 23, 2019, https://www.justice.gov/opa/speech/attorney-general-william-p-barr-delivers-keynote-address-international-conference-cyber.

11. Digiconomist, "Bitcoin Energy Consumption Index," accessed September 16, 2023, https://digiconomist.net/bitcoin-energy-consumption.

12. Herbert Lin, "A Retrospective Post-Quantum Policy Problem," Lawfare, September 14, 2022, https://www.lawfaremedia.org/article/retrospective-post-quantum-policy-problem.

13. Federal Reserve, "Central Bank Digital Currency (CBDC): Frequently Asked Questions," accessed August 15, 2023, https://www.federalreserve.gov/cbdc-faqs.htm.

14. Darrell Duffie and Elizabeth Economy, eds., *Digital Currencies: The US, China, and the World at a Crossroads*, (Stanford, CA: Hoover Institution, 2022), https://www.hoover.org/sites/default/files/research/docs/duffie-economy_digitalcurrencies_web_revised.pdf.



04

MATERIALS SCIENCE

KEY TAKEAWAYS

- Materials science is a foundational technology that underlies advances in many other fields, including robotics, space, energy, and synthetic biology.
- Materials science will exploit AI as another promising tool to predict new materials with new properties and identify novel uses for known materials.
- The structure of funding in materials science does not effectively enable transition from innovation to implementation. Materials-based technology that has been thoroughly tested at the bench scale may be too mature to qualify for basic research funding (because the high-level basic science is understood) but not mature enough to be directly commercialized by companies.

Overview

Materials are everywhere, from macro features that are visible to the naked eye to microscopic features thousands of times smaller than the diameter of a single human hair. They shape the objects of everyday life and give rise to new possibilities. Materials science cuts across technological areas, contributing to everything from the development of stronger and lighter materials for aircraft to more efficient and less heavy solar cells, better semiconductors, biocompatible materials for medical implants, more stable electrodes for batteries, and easily manufactured and recyclable plastics.

The goal of materials science is to understand how the structure of a material influences its properties and how processing the material can change its structure and therefore its performance. This knowledge can then be used to design new materials with desirable properties for specific uses. The ultimate aspiration, which remains a long way off, is to be able to create materials on demand by specification—put in a request for a material with properties X, Y, and Z, and a 3-D printer produces it for you.

Broadly speaking, materials science and engineering research focuses on four major areas. The first is characterizing the properties of materials. The second is modeling materials, which involves predicting material properties based on atomic principles. The third is synthesizing materials with precise control to verify whether their properties are as predicted. The fourth area is manufacturing and processing materials with well-characterized properties in sufficient quantities for practical applications.

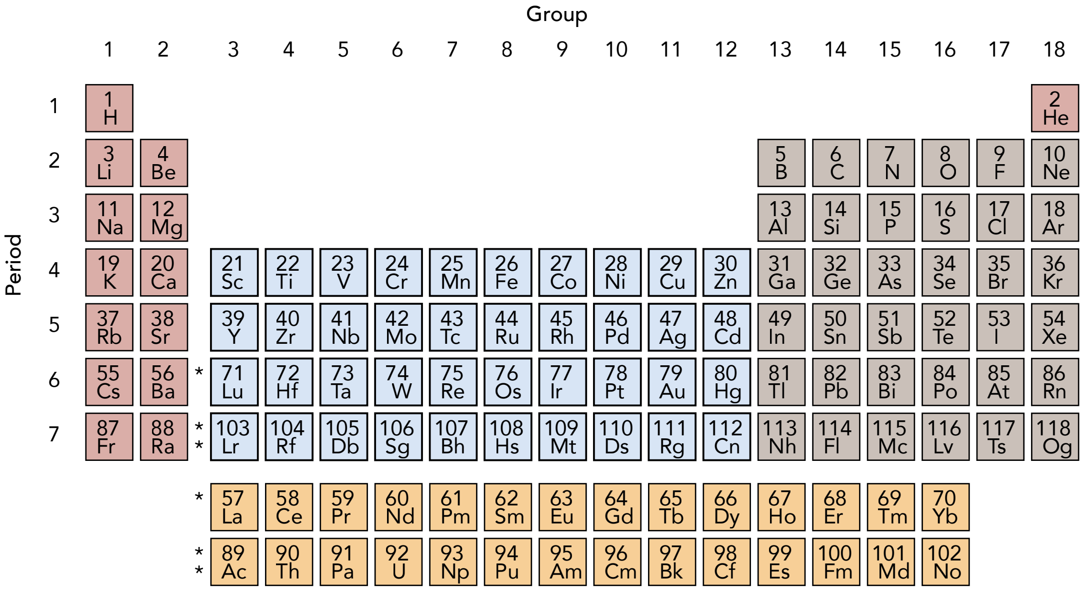
Basics of Materials Science

All materials are composed of atoms. The periodic table of the elements (figure 4.1) lists all the known types of atoms. Certain atoms can be combined into molecules that have vastly different properties than the atoms alone. For example, table salt consists of sodium and chlorine, which are elements. Sodium burns on contact with water, chlorine is a poisonous gas, and yet the table salt we consume every day is a completely different substance.

There are two important points to note about the periodic table. First, there are a lot of elements—ninety-two naturally occurring ones and twenty-six that can be observed only in laboratory conditions. That's a lot of building blocks from which different materials and molecules can be synthesized, and in fact, an astronomically large number of different compounds are possible. The challenge for materials science is to sift through this vast array of possibilities to find the ones that are useful.

The second important point is that the elements in the periodic table are lined up in a certain order. Elements in the same column have properties that are often similar in key ways. This means that insights developed through experimentation or calculation on one element can be transferred, with modifications, to another element above or below it in the periodic table.

FIGURE 4.1 Periodic table of the elements

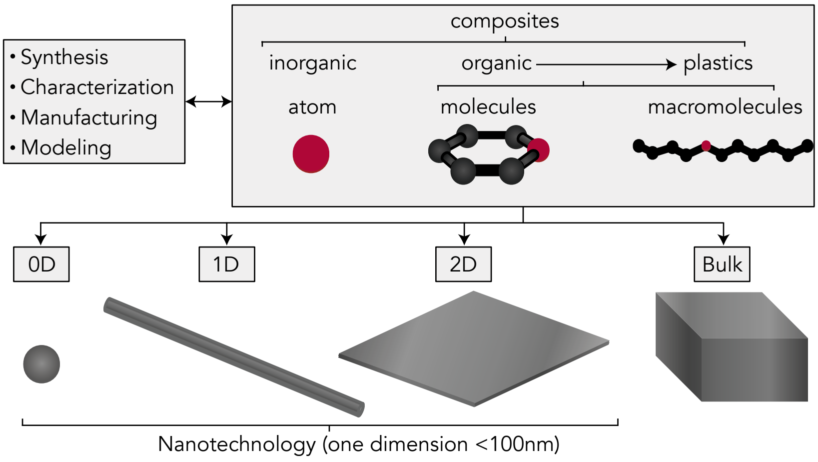


Group

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: Wikimedia Commons, CC BY-SA 4.0

FIGURE 4.2 The basic layout of materials science



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.

Molecules, in turn, can be linked together into structures called macromolecules (see figure 4.2). These can occur naturally, such as proteins, DNAs, and cellulose, or can be synthesized artificially, resulting in polymers/plastics, for example. Plastics are particularly useful because the long chains of macromolecules are often more flexible. Research on new macromolecular structures can be used to develop plastic materials that are easier to recycle or that hold advantageous mechanical properties while weighing less than metals.

Key Developments

Present-Day Applications

Some interesting applications from studying materials science include:

Biomedical applications Wearable electronic devices made from flexible materials conform to skin or tissues and serve specific sensing or actuating functions. More specifically, wearable electronic devices or "e-skin" can sense external stimuli such as temperature and pressure and encode these stimuli into electrical signals.[1] For example, a "smart bandage" with integrated sensors and simulators can accelerate healing of chronic wounds by 25 percent.[2]

Novel and recyclable plastics Researchers are developing new sustainable methods to couple molecules into polymers for deconstructable plastics that are easier to recycle.[3] New electrically conductive polymers are also a focus of study. Electrical conductivity in flexible materials such as plastics can be achieved by inserting specific bonds between individual atoms that make up the material backbone. This allows for the fabrication of flexible electronic devices such as wearable sensors and foldable screens for mobile devices.

Energy materials Materials design and processing is integral to decarbonization efforts through the electrification of transportation and industry. Some challenges persist, however, including storing energy from intermittent energy sources, such as solar and

> **The ultimate aspiration . . . is to be able to create materials on demand by specification.**

wind, in batteries. Therefore, designing batteries with materials and architectures that enable quick recharging and long stability while reducing costs will be crucial. Important discoveries in engineering battery electrode materials have been made.[4] Studying the electrolyte-electrode interface in batteries has also led to higher performing and more stable electrolytes in batteries.[5]

Additive Manufacturing

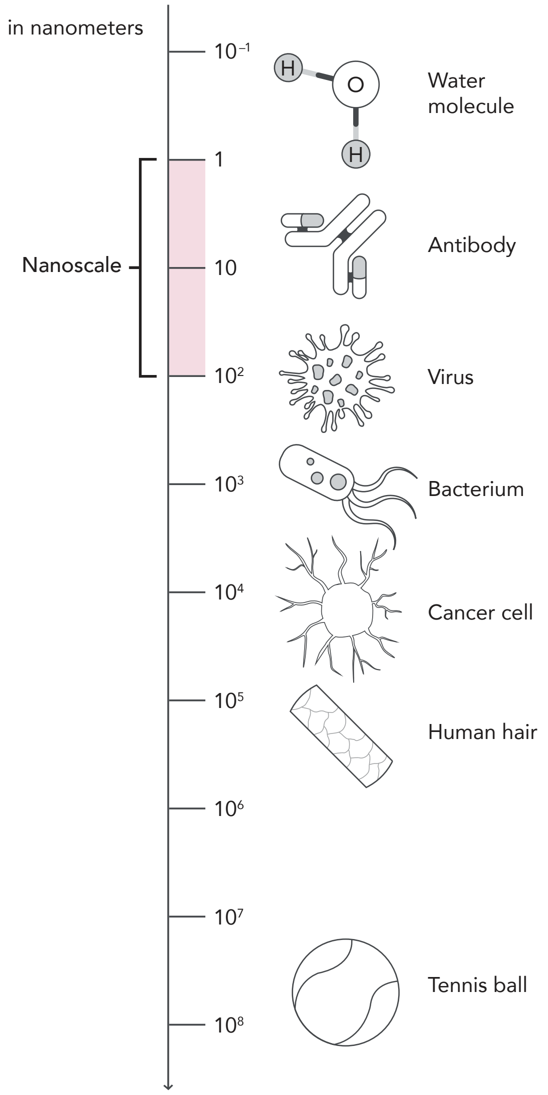
One promising advance in materials processing over the past fifteen years is additive manufacturing, or 3-D printing. A novel method termed continuous liquid interface production (CLIP) has been established that uses directed ultraviolet light to pattern structures from a polymer resin.[6]

This technology has been used to make customized football helmet liners,[7] and a number of companies have sprung up to commercialize and scale up additive manufacturing both by producing stand-alone products and by collaborating with multinational companies. More recent active research in 3-D printing includes scaling down 3-D printable feature sizes and exploring methods to 3-D print with conductive materials and artifacts using multiple materials at once.

Nanotechnology

Nanotechnology is a large and growing subfield of materials science. Size has a profound impact on the properties of a material. Figure 4.3 shows different length scales compared to a water molecule (which is below a nanometer), a human hair (roughly

FIGURE 4.3 The size of nanoscale objects



10^5 nanometers), and a tennis ball (at 10^8 nanometers). A structure is typically referred to as nanoscale if at least one dimension is in the 1–100 nanometer range.

In the past twenty years, nanoscience and nanotechnology have attracted enormous interest, for two reasons. First, many significant biological organisms (such as viruses and proteins) are nanoscale in size. Second, it turns out that 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.[8] Materials that are smaller than about 100 nanometers in one dimension, two dimensions, or in all dimensions are called nanosheets, nanowires, and nanoparticles, respectively.

Quantum dots—for which the Nobel Prize in Chemistry was awarded in 2023—have garnered public attention through their use in televisions. Quantum dots are metallic, carbonaceous, or semiconductor spherical nanocrystals that emit bright monochromatic light in response to excitation by a light source with a higher energy, such as blue light from the back panel in a display.[9] Quantum dots are a model example of variable material properties due to scale as their optoelectronic properties differ from those of the same bulk material. The diameter of quantum dots shifts the color of light that they emit, with larger quantum dots emitting longer wavelengths. This allows for tunable light emission based on the desired application.

Some current applications of quantum dots include:

Medical imaging Quantum dots are being used to improve the contrast of biomedical imaging, for example, as in fluorescent markers to allow selective labeling of biological structures in vitro and in vivo.[10] Additionally, biocompatible nanomaterials can be employed as optical probes to sense mechanical forces and electrical fields in biological organisms, thus circumventing specialized and bulky equipment, opening the possibility of new experiments.[11]

Solar cells Quantum dots can improve the efficiency of solar cells. Their ability to absorb different frequencies of light means they can potentially capture more of the solar spectrum, boosting the performance of solar panels.[12]

Sensors Quantum dots and plasmonic nanoparticles can be used in sensors for detecting chemicals and biological substances.[13]

Anticounterfeiting Quantum dots can be embedded in labels to defend against counterfeiting[14.]

Some examples of applications of other nanomaterials include:

Pharmaceutical delivery An injectable polymer-nanoparticle hydrogel, for example, was developed so the delivery of drugs, proteins, and cells can be precisely controlled, enabling months-long release of entrapped cargo.[15] The efficacy of insulin administration can also be improved through this research.[16] Nanoparticles can be engineered to permeate the blood–brain barrier, delivering drugs to treat neurodegenerative diseases.[17]

Vaccine stabilization Nanoassemblies can be used to stabilize certain types of vaccines, notably mRNA vaccines, by encapsulating them.[18] In this form, it is easier to inject the vaccine into the human body and to release it over time inside the body in a controlled manner.

Smart windows Silver nanowires arranged into a thin film on a window become a transparent conductive film rather than the familiar reflective mirror from silver behind the window. Running a current through the film can then change the opacity of the window electrically.[19]

2-D semiconductors, graphene, carbon nanotubes, and nanoscale materials These are at the forefront of the next generation of high-tech electronic devices. Active research efforts are dedicated to designing new methods to integrate 2-D or carbon nanotube semiconductors into electronics that are currently silicon based to increase their energy efficiency and heat management.[20]

Higher-capacity batteries High-performance lithium battery anodes have been developed by integrating silicon nanowires as an anode material. When bulk silicon is used as an anode, it undergoes significant changes in volume as the battery charges and discharges, often leading to mechanical failure. Use of silicon nanowires bypasses this problem and increases battery capacity by a factor of ten.[21]

Catalysis Catalysts are used to accelerate chemical reactions, and nanomaterials are well suited for this role.[22] Nanoparticles are particularly well suited for this task, as they contain a high number of active sites per unit mass and can be chemically architected to catalyze various chemical reactions. Advances have been made in converting CO_2 to value-added chemicals using electrified nanoparticle catalysts and in employing palladium catalysts for the combustion of methane, which could improve the efficiency of electricity generation from methane.[23] Nanocatalysts have also been used to improve the rate at which hydrogen can be produced from water through electrolysis.[24] The challenges include developing catalysts that are sufficiently active, stable, and low in cost to produce hydrogen in large quantities and inexpensively.[25]

Over the Horizon

Impact of New Technologies

LOW-CARBON STEEL AND CEMENT PRODUCTION

As an example of how materials science could have impact on a large scale, note that steel and concrete are critical building materials. World production of concrete is some 30 billion tons per year. For comparison, the weight of all the concrete in New York City is around 750 million tons, according to the US Geological Survey.[26] The Hoover Dam involves about 10 million tons of concrete.

Cement production is an extremely carbon-intensive activity, contributing to 8 percent of CO_2 emissions. Limestone is burned to produce lime, thereby releasing CO_2. A number of approaches have potential for reducing the CO_2 footprint of cement production. One focuses on using different material inputs in the production process that release less CO_2. These inputs are the basis for "supplementary cementitious materials," which are formulated differently than traditional Portland cement but nevertheless can substitute for Portland cement in many cases. Another approach incorporates captured CO_2 into concrete during the curing process.[27]

These techniques are all well proven, but further research is needed to make them economically competitive with traditional CO_2-intensive methods of production.

THE APPLICATION OF AI TO MATERIALS SCIENCE

An interesting topic today is whether AI machine learning and modeling will be useful in predicting properties of new materials based on what is known about existing materials.[28] Success has been seen with less complicated materials, but much is to be done and more data are needed for complex materials.

Challenges of Innovation and Implementation

The 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 feasibility of potential large-scale manufacturing. At this point, the technology is too mature to qualify for most research funding—because the basic science questions do not address issues related to scaling up—but not mature enough to be commercialized into actual companies. Neither government nor venture capital investors are particularly enthusiastic about funding pilot projects, so different forms of funding are required to bridge this gap between bench-scale research and company-level investment. The support could even go one step further and establish national rapid prototyping centers, where academic researchers find the help and tools necessary to build prototypes and pilot plants for their technology.

Research processes born in the past are also ill suited to the rapid transitions to real-world application. 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 society challenges calls for a more scalable system-level approach that involves extensive rapid prototyping and reliable demonstrations to provide feedback on and fill in gaps of knowledge.

Current infrastructure makes 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. With typically less than a thirty-minute window to place a device on a patient and gather data, any malfunction, such as a sudden equipment failure or a loose wire, can jeopardize the entire experiment and potentially halt future patient interactions. The laboratory-assembled devices may not meet this standard of reliability, even if they do demonstrate the value of the underlying science.

Policy, Legal, and Regulatory Issues

REGULATION OF PRODUCTS INCORPORATING NANOMATERIALS

As with regulation in other areas of technology, concerns arise about the appropriate balance between promoting public safety from possible downside risks and the imperatives of innovation to move quickly and leapfrog possible competitors. In the biomedical space, the FDA created a Nanotechnology Regulatory Science Research Plan in 2013.[29] Today, FDA regulation and review of nanotechnology is governed by Executive Order 13563.[30] Outside of biomedicine, regulation and infrastructure for nanomaterials research from the government side is largely based in agencies of the National Nanotechnology Initiative, which include the Department of Energy, the National Cancer Institute, the National Institutes of Health (NIH) more broadly, the National Institute for Standards and Technology (NIST) in the Department of Commerce, and the National Science Foundation (NSF).

TOXICITY AND ENVIRONMENTAL ISSUES

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 and could affect biological systems in harmful ways. Nanoscale particles inhaled into the lungs, for example, may lodge themselves permanently, causing severe health outcomes, including pulmonary inflammation, lung cancer, and penetration into the brain and skin.[31]

Furthermore, because engineered nanoparticles are, by definition, new to the natural environment, they pose unknown dangers to humans and the environment. There are concerns about incorporating nanomaterials into products that enter that environment at the end of their life cycles. As nanomaterials are employed in and considered for electronic and energy products, it is paramount that those materials safely degrade or can be recycled at the end of a product's life. Policy will be particularly important in shaping responsible end-of-life solutions for products incorporating nanomaterials.

FOREIGN COLLABORATION AND COMPETITION

Historically, the United States has led the world in nanotechnology, but the gap between the United

States and China has narrowed. Notably, in 2016, the president of the Chinese Academy of Sciences openly announced Beijing's ambition to compete in the field of nanotechnology.[32]

As great power competition intensifies, many researchers are concerned that fundamental research could now be considered export controlled. Policy ambiguity can inadvertently hinder innovation by creating obstacles for non-US researchers wishing to contribute to work in the United States and by deterring international collaborations, allies, and partners who are important for advancing the field. In nanomaterials, for example, researchers in Korea are making significant strides with biomedical applications and consumer electronics. There is an urgent need for clarification of these policies, particularly delineating fundamental research and export-controlled research.

NOTES

1. Weichen Wang 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 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 D. Feist, Daniel C. Lee, and Yan Xia, "A Versatile Approach for the Synthesis of Degradable Polymers Via Controlled Ring-Opening Metathesis Copolymerization," *Nature* Chemistry 14 (2022): 53–58, https://doi.org/10.1038/s41557-021-00810-2.

4. Dingchang Lin, Yayuan Liu, and Yi Cui, "Reviving the Lithium Metal Anode for High-Energy Batteries," *Nature* Nanotechnology 12 (2017): 194–206, https://doi.org/10.1038/nnano.2017.16.

5. Zhiao Yu et al., "Rational Solvent Molecule Tuning for High-Performance Lithium Metal Battery Electrolytes," *Nature Energy* 7 (2022): 94–106, https://doi.org/10.1038/s41560-021-00962-y.

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

7. Benjamin Perez, "Riddell's 3D Printed Helmet Liners Are the MVPs of the NFL Playoffs," *3DPrint.com*, January 24, 2023, https://3dprint.com/297161/riddells-3d-printed-helmet-liners-are-the-mvps-of-the-nfl-playoffs.

8. Hui Pan and Yuan Ping Feng, "Semiconductor Nanowires and Nanotubes: Effects of Size and Surface-to-Volume Ratio," ACS Nano 2, no. 11 (November 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 (November 2006): 1107–10, https://doi.org/10.1126/science.1130557.

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

10. X. Michalet et al., "Quantum Dots for Live Cells, in Vivo Imaging, and Diagnostics," *Science* 307, no. 5709 (January 2005): 538–44, https://doi.org/10.1126/science.1104274.

11. Randy D. Mehlenbacher et al., "Nanomaterials for In Vivo Imaging of Mechanical Forces and Electrical Fields," *Nature Reviews Materials* 3, no. 17080 (2018), https://doi.org/10.1038/natrevmats.2017.80.

12. Prashant V. Kamat, "Quantum Dot Solar Cells: Semiconductor Nanocrystals as Light Harvesters," *Journal of Physical Chemistry* C 112, no. 48 (October 2008): 18737–53, https://doi.org/10.1021/jp806791s; Ralph Nuzzo et al., "Light Material Interactions in Energy Conversion (Final Report)," Office of Scientific and Technical Information, US Department of Energy, April 1, 2019, https://doi.org/10.2172/1504275.

13. Babatunde Ogunlade et al., "Predicting Tuberculosis Drug Resistance with Machine Learning-Assisted Raman Spectroscopy," arXiv, Cornell University, June 9, 2023, https://doi.org/10.48550/arXiv.2306.05653; Fareeha Safir et al., "Combining Acoustic Bioprinting with AI-Assisted Raman Spectroscopy for High-Throughput Identification of Bacteria in Blood," *Nano Letters* 23, no. 6 (March 2023): 2065–73, https://doi.org/10.1021/acs.nanolett.2c03015.

14. Yang Liu et al., "Inkjet-Printed Unclonable Quantum Dot Fluorescent Anti-Counterfeiting Labels with Artificial Intelligence Authentication," *Nature Communications* 10, no. 2409 (2019), https://www.nature.com/articles/s41467-019-10406-7.

15. Hector Lopez Hernandez et al., "Non-Newtonian Polymer–Nanoparticle Hydrogels Enhance Cell Viability During Injection," *Macromolecular Bioscience* 19, no. 1 (January 2019), https://doi.org/10.1002/mabi.201800275.

16. Caitlin L. Maikawa et al., "Formulation Excipients and Their Role in Insulin Stability and Association State in Formulation," *Pharmaceutical Research* 39 (2022): 2721–28, https://doi.org/10.1007/s11095-022-03367-y; Joseph L. Mann et al., "An Ultrafast Insulin Formulation Enabled by High-Throughput Screening of Engineered Polymeric Excipients," *Science Translational Medicine* 12, no. 550 (July 2020), https://doi.org/10.1126/scitranslmed.aba6676.

17. Vladimir P. Torchilin, "Multifunctional, Stimuli-Sensitive Nanoparticulate Systems for Drug Delivery," *Nature Reviews Drug Discovery* 13 (November 2014): 813–27, https://doi.org/10.1038/nrd4333; Cláudia Saraiva et al., "Nanoparticle-Mediated Brain Drug Delivery: Overcoming Blood-Brain Barrier to Treat Neurodegenerative Diseases," *Journal of Controlled Release* 235 (August 2016): 34–47, https://doi.org/10.1016/j.jconrel.2016.05.044.

18. Norbert Pardi et al., "mRNA Vaccines—A New Era in Vaccinology," *Nature Reviews Drug Discovery* 17 (2018): 261–79, https://doi.org/10.1038/nrd.2017.243.

19. Zhiqiang Niu et al., "Synthesis of Silver Nanowires with Reduced Diameters Using Benzoin-Derived Radicals to Make

Transparent Conductors with High Transparency and Low Haze," *Nano Letters* 18, no. 8 (2018): 5329–44, https://doi.org/10.1021/acs.nanolett.8b02479.

20. Weisheng Li 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.

21. Candace K. Chan 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.

22. U. P. M. Ashik et al., "Chapter 3. Nanomaterials as Catalysts," in *Applications of Nanomaterials: Advances and Key Technologies*, ed. Sneha Mohan Bhagyaraj et al. (Cambridge, MA: Elsevier, 2018), 4582, https://doi.org/10.1016/B978-0-08-101971-9.00003-X.

23. Weixin Huang et al., "Steam-Created Grain Boundaries for Methane C–H Activation in Palladium Catalysts," *Science* 373, no. 6562 (September 2021): 1518–23, https://doi.org/10.1126/science.abj5291; Chengshuang Zhou et al., "Steering CO_2 Hydrogenation Toward C–C Coupling to Hydrocarbons Using Porous Organic Polymer/Metal Interfaces," *Proceedings of the National Academy of Sciences* 119, no. 7 (February 2022), https://doi.org/10.1073/pnas.2114768119.

24. Thomas F. Jaramillo et al., "Identification of Active Edge Sites for Electrochemical H_2 Evolution from MoS_2 Nanocatalysts," *Science* 317, no. 5834 (July 2007): 100–102, https://doi.org/10.1126/science.1141483.

25. Zhi Wei Seh et al., "Combining Theory and Experiment in Electrocatalysis: Insights into Materials Design," *Science* 355, no. 6321 (January 2017), https://doi.org/10.1126/science.aad4998.

26. Tom Ough, "New York's Skyscrapers Are Causing It to Sink—What Can Be Done?," BBC, May 23, 2023, https://www.bbc.com/future/article/20230523-new-yorks-skyscrapers-are-causing-it-to-sink-what-can-be-done-about-it.

27. Liang Li and Min Wu, "An Overview of Utilizing CO2 for Accelerated Carbonation Treatment in the Concrete Industry," *Journal of CO2 Utilization* 60, no. 10200 (2022), https://doi.org/10.1016/j.jcou.2022.102000.

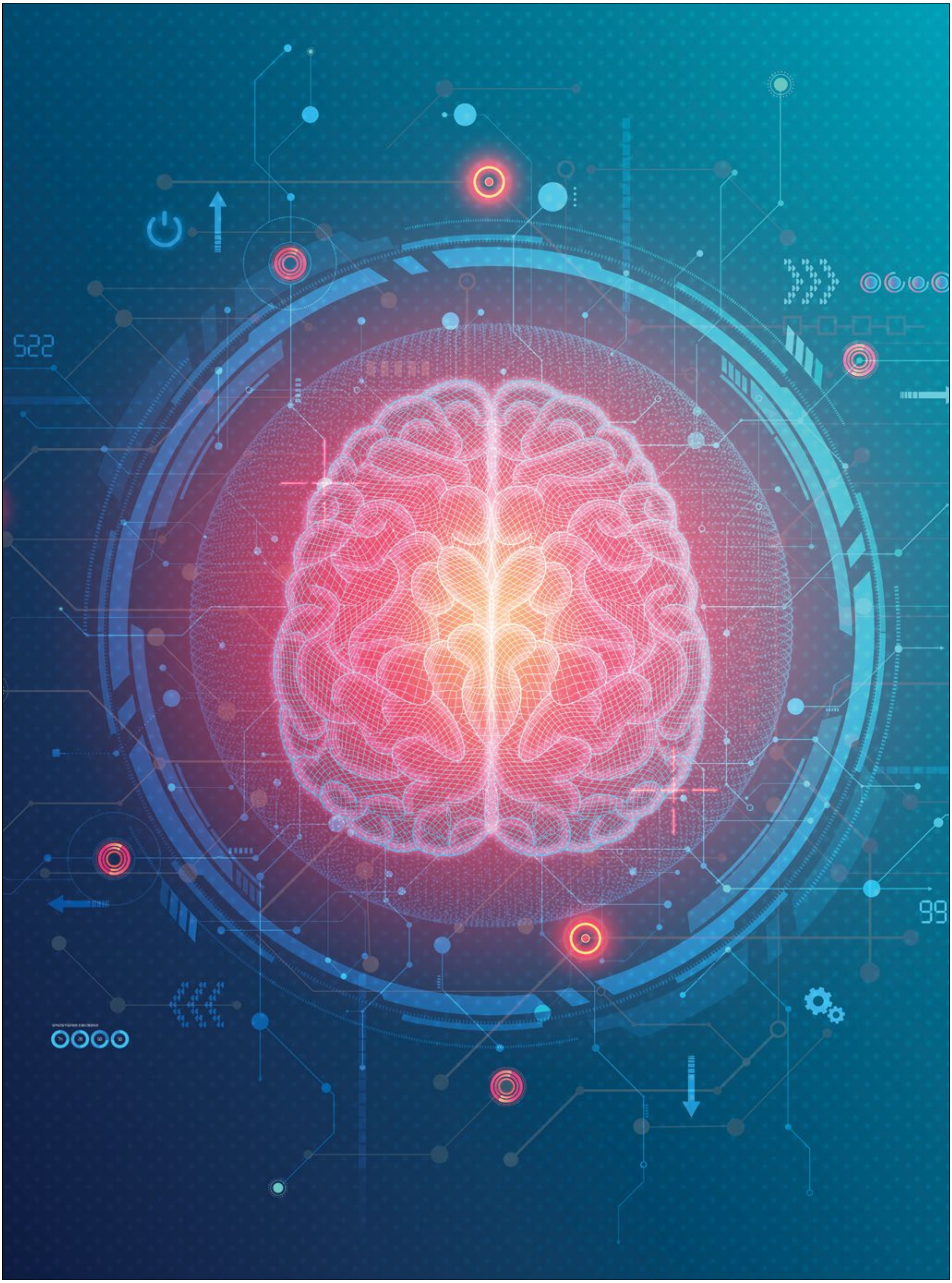
28. Steven G. Louie et al., "Discovering and Understanding Materials through Computation," *Nature Materials* 20 (2021): 728–35, https://doi.org/10.1038/s41563-021-01015-1.

29. 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.

30. 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.

31. 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.

32. Chunli Bai, "Ascent of Nanoscience in China," *Science* 309, no. 5731 (July 2005): 61–63, https://doi.org/10.1126/science.1115172.


522
99

05

NEUROSCIENCE

KEY TAKEAWAYS

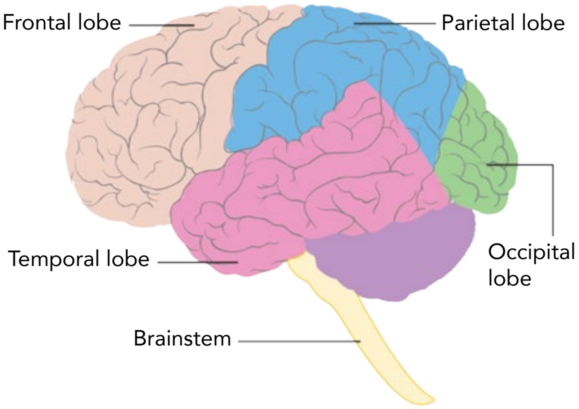
- Popular interest in neuroscience vastly exceeds the actual current scientific understanding of the brain, giving rise to overhyped claims in the public domain that revolutionary advances are just around the corner.
- Advances in computing have led to progress in several areas, including understanding and treating addiction and neurodegenerative diseases, and designing brain-machine interfaces.
- American leadership is essential for establishing and upholding global norms about ethics and human subjects research in neuroscience.

Overview

Neuroscience is the study of the human brain and the nervous system, their structure and function, healthy and diseased states, and life cycle from embryonic development to degeneration in later years.[1] Today's product applications of science support a growing market. Already a $32 billion market in 2021, the market for such products based on neuroscience is forecast to grow nearly 4 percent annually this decade to $41 billion, driven by increasing cases of neurological disorders like Parkinson's and Alzheimer's.[2]

The human brain (see figure 5.1) 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.[3] 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.

FIGURE 5.1 The human brain



Source: Cancer Research UK / Wikimedia Commons, CC BY-SA 4.0

The brain contains some one hundred billion nerve cells, or neurons, which are the fundamental building blocks of the brain and nervous system. These cells sense the physical world, transmit information to the brain, process information, and send information from the brain to other parts of the body. The physical feature neurons use to connect to other neurons is called the synapse. Each neuron can have just a few or a hundred thousand synapses, though on average each neuron has several thousand synapses. Synapses are the structures that mediate most communication between neurons. The upstream neuron produces a chemical that is expelled at the synapse, diffuses for a very short distance, and finally is sensed by a receptor molecule on the downstream neuron. The prevalent view in the field is that memory is stored in the network of synapses.

Neuroscience covers a wide range of subfields: the nervous system's bodily structure (neuroanatomy), chemicals that modulate the nervous system (neurochemistry), nervous system functions (neurophysiology), the role of the nervous system in actions (behavioral neuroscience), and the role of the nervous system in thoughts (cognitive neuroscience). Many practical applications could benefit from neuroscience research, including the development of treatments for neurological and psychiatric disorders such as epilepsy, learning disabilities, cerebral palsy, and anxiety, as well as Alzheimer's disease and other neurodegenerative disorders.

Key Developments

This report 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). The most mature of these is the first.

Neuroengineering and the Development of Brain-Machine Interfaces

A brain-machine interface is a device that maps neural impulses from the brain to a computer, and vice versa. The potential applications for mature brain-machine interface technologies are wide ranging: sensory replacement or augmentation, limb replacement, direct mind-to-computer interfacing, and even computer-assisted memory recall and cognition are all within the theoretical realms of possibility. However, despite compelling headlines about mind-reading chip implants, that is still mostly science fiction. Even with tremendous interest and rapid progress in neuroscience and engineering, there are exceptionally few areas of the brain for which we have the necessary theoretical understanding of how neurocircuits work, and we also have not solved the technical problems of safely implanting electrodes in the brain.

An 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

> **There are exceptionally few areas of the brain for which we have the necessary theoretical understanding of how neurocircuits work.**

diseases are blind because their light-detecting cells do not work. The 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 from those images, therefore bypassing the nonfunctional light-detecting cells.[4]

The theoretical side of this project—the science—involves recording spontaneous neural activity to identify cell types and their normal signals, identifying how electrodes activate cells, and understanding how to stimulate retinal ganglion cells to represent an image so that this information can be transmitted by the optic nerve. Solving the technical problems calls for significant engineering know-how in translating the scientific understanding of the stimulation algorithm into practical applications, experimental recordings, and fabrication and packaging of the electrode into the device—and in surgical techniques.

This effort requires a multidisciplinary team of neuroscientists, ophthalmologists, and surgeons working with electrical engineers and computer scientists.

The artificial retina project is the most mature brain-machine interface to date. The retina, a part of the central nervous system, is well suited as an experimental environment, as its stimuli (light) is experimentally controllable and can be captured by a digital camera. The retina 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 interfaces 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 limb prosthetic. Here, feeding high-dimensional patterns of neural activity into an AI algorithm suffices to control an artificial limb without requiring direct control of neural functions, a control that remains beyond our current scientific understanding. Another example of a unidirectional interface is the demonstrated use of data from functional magnetic resonance imaging (fMRI) studies of an individual to train a computer to reconstruct thoughts formulated as language from other fMRI data obtained in real time from that individual[5] and to measure emotional responses to informational stimuli in real time.[6]

Such demonstrations hint at the prospect of other brain-machine interfaces in the future, such as computer-assisted memory recall, even if the suite of specific future applications is unclear. The scope and feasibility of such applications will be determined by advances in neuroscientific theory and in technical solutions to engineering problems, such as probe density, spread, and penetration into deep-layer tissues.

Over the Horizon

The Impact of Neuroscience

NEUROHEALTH AND NEURODEGENERATION

Neurodegeneration is a major challenge as humans live longer. In the United States alone, the annual cost of Alzheimer's treatment is projected to explode from $305 billion today to $1 trillion by 2050.[7] Diseases like Alzheimer's and Parkinson's surge in frequency with age—while just 5 percent of 65–74-year-olds have Alzheimer's, this rises to 33 percent of those over 85.[8] The percentage of adults with Parkinson's disease demonstrates a similar rising frequency with age.[9] As modern medicine and society enable longer life spans, the human body and brain remain maladapted to maintaining nervous system function for decades past childbearing age.

Effective treatments for neurodegenerative disorders such as Alzheimer's are still far from sufficient despite decades of research. Alzheimer's disease occurs when two different types of proteins in the brain fold improperly, which eventually leads to neuron death. Only in the last two years have drugs aimed at clearing one of these proteins been approved by the FDA, albeit with limited though real therapeutic benefit and significant side effects. These drugs have also been subject to significant controversy. One drug, aducanumab, received approval in 2021 in the United States but not the European Union. A scientific advisory panel at the FDA voted against approval, citing minimal therapeutic benefits and high risk of complications. The FDA overruled the advisory panel, which led to three of the nine members resigning in protest.[10] A second drug, lencanemab, was approved in 2023, but again detractors suggest the treatment may not be worth the risks.[11] Still, these are the first drugs to suggest that the slowing of neurological disease progression is possible.

While current treatments for Alzheimer's disease are less effective than would be desired, there is reason for guarded optimism in the coming years. Gene therapy approaches targeting other proteins associated with Alzheimer's have recently entered clinical trials. Powerful diagnostic tools such as tau and amyloid PET scans, identification of biomarkers, and identification of genetic risk factors allow for increasing early detection and diagnosis, which might make it easier to fight the disease. Advances in personalized medicine also leave researchers and clinicians hopeful.

Another form of neurodegeneration results from traumatic brain injury (TBI), which can manifest itself in a range of complex symptoms and pathologies.[12] 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 predominantly exhibit 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.

NEURODISCOVERY AND THE SCIENCE OF ADDICTION

Researchers are working to understand the neural basis of addiction and of chronic pain while working with psychiatrists and policymakers to address the opioid epidemic.[13] The economic costs of the opioid epidemic are difficult to calculate, but estimates range from $100 billion to $1 trillion a year when the loss of lifetime earnings of overdose victims is included.[14] The number of opioid deaths in the United States has risen sharply over the last ten to fifteen years, from 21,000 in 2010 to 80,000 in 2021,[15] which places opioid overdose as one of the top ten leading causes of death in the United States, comparable to diabetes and Alzheimer's disease.[16]

Economic, societal, and political factors all have a role in the epidemic. But neuroscience has a potentially important role to play as well. For example, a nonaddictive painkiller drug as effective as current-generation opioids would be transformative,[17] but detractors note that relief from pain itself is pleasurable and thus may be behaviorally addictive as well. Indeed, it is possible to become addicted to behaviors that do not involve consumption of a drug—consider gambling, sex, or technology addictions. Heroin and oxycontin themselves were famously initially marketed as nonaddictive alternatives to painkillers of the day.[18]

Though safer or less addictive painkillers would help reduce the burden of the opioid epidemic, other approaches are relevant to neuroscience, such as reducing the need for opioids or aiding in recovery from addiction. Consider the problem of relapse in tackling addiction. Particularly relevant for opioid use, scientists have found that the brain mechanisms leading to an initial opioid addiction differ significantly from those that trigger a relapse.

Neuroscientists may be able to assist in the social aspects of recovery from opioid addiction. It turns out that opioid receptors are found in many areas of the brain and affect diverse functions, including neural circuits related to the desire for social interaction. When an individual goes into opioid withdrawal, these areas of the brain are affected—and the individual often develops an aversion to social interactions. Such an aversion is a significant challenge to recovery since social interactions are often key to helping an individual to cope with the vulnerabilities associated with recovery. Essentially, their brains miscalculate the rewarding value of human connection, undermining their recovery process.

Stanford neuroscientists have recently identified a neurological pathway that is responsible for the onset of this social aversion.[19] If this study conducted in mice generalizes to humans, it may be possible to develop drugs that inhibit social aversion during withdrawal and thereby assist patients in seeking help or companionship from friends, families, recovery programs, and doctors.

Finally, it is widely recognized that chronic pain is a driver of opioid misuse.[20] Chronic pain is a widespread condition—an estimated 20 percent of adults in the United States experience chronic pain and around 7 percent have intense chronic pain that results in substantial impacts on daily life.[21] Compared to other alternatives, prescription opioids are unparalleled for managing acute pain, but rapid onset of tolerance and their addictive properties make them unsuitable for long-term use.

But what if it were possible to block the induction of chronic pain entirely? Soldiers with severe injuries, including compound fractures and open wounds, don't always report immediate pain. This fact suggests that pain isn't just governed by ascending signals—from the injury site to the brain—but also that the central nervous system can exert influence. That is, our brain can control whether we sense pain, a phenomenon known as descending pain control.[22] Certain neurons act as switches that control whether pain signals from the injured site reach the cortex. This is relevant to opioid use because opioid receptors are found in these neurons, and opioids inhibit pain by stimulating the neurons that block the ascending pain signals.[23]

On the other hand, opioids are addictive substances. Drug addiction is a compulsive use of a drug despite its long-term negative consequences. Regardless of the specific drug in question, the mechanism of drug addiction operates in the same way: consumption of the drug releases excessive amounts of dopamine in the brain, which then goes to the nucleus accumbens that are involved in reward and finally to the prefrontal cortex, which controls executive functions like goal selection and decision making. Similar mechanisms also appear to operate in behavioral addictions, such as addiction to gambling, technology, or video games.[24]

This reward mechanism evolved to reinforce activities that are crucial to survival, such as foraging for food and procreation. The release of dopamine serves as a robust positive signal, strengthening and reinforcing the activity that led to its release in the first place. But drugs appear to hijack this reward system, causing a surge in dopamine that far surpasses what the natural system can produce. This creates a potent lure that can make overcoming drug addiction particularly challenging, as the drugs tap into and significantly amplify this natural reinforcement system.

SCIENTIFIC THEORETICAL AND TECHNICAL ENGINEERING CHALLENGES

Contrasting the work on the artificial retina and the work on the science of neurodegeneration and addiction illustrates the dual-pronged nature of neuroscience applications. They have two primary components: a scientific component that focuses on identifying relevant brain circuits and understanding how they function and compute, and a technical engineering component that focuses on how to safely stimulate the relevant brain circuits to create the desired responses.

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 over industry. Certainly, there are research programs in industry that solve basic biological questions in neuroscience, but these are necessarily and economically 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. 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 application by aiding in high-throughput screening—the use of automated equipment to rapidly test samples—and 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, technical engineering, and final application.

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 expected to continue to grow considerably over the coming years. The Society for Neuroscience's annual meeting draws close to thirty thousand attendees.[25]

Science fiction and fantastical headlines fuel beliefs that mind-reading technology, brains controlled by computers, and other dystopias are imminent. In reality, comprehending the brain's staggering complexity remains in its early stages. Most advances involve incremental progress expanding our theoretical foundations rather than 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 leave many open to dubious proclamations or pseudoscience.

Policy, Legal, and Regulatory Concerns

DRUG POLICY AND NEUROSCIENCE RESEARCH

The Controlled Substances Act (CSA) of 1970 as amended governs US policy regarding regulation of the manufacture, importation, possession, use, and distribution of certain substances. Substances on Schedule I are drugs or other substances with a high potential for abuse and not "currently accepted" for medical use in the United States. No research exceptions are provided for Schedule 1 substances such as cannabis or MDMA (often known as Ecstasy or Molly), which have potential for medical use that might be realized through research. (At the time of this writing, the Biden administration is reportedly considering the reassignment of marijuana to Schedule 3, a schedule with fewer restrictions.[26]) Placement of drugs on Schedule 1 sharply constrains researchers because these potentially helpful drugs are difficult to obtain. This constraint also denies the public the benefits that might flow from such research—such as better medical treatments—and potentially harms the public if, for example, individual states legalize certain drugs without adequate research into their safety, addictiveness, and public health impacts.

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. Thus, under a 2005 US Supreme Court ruling, minors under eighteen years of age cannot be subject to the death penalty for crimes they committed because adolescent brains are not fully developed, putting minors at higher risk of impulsive, irrational thoughts and behaviors.[27]

THOUGHT IMPLANTS

The possibility that information can be implanted directly into a person's consciousness is a potential future problem. As government is still figuring out how to regulate internet forums that influence what people believe and how they feel—a problem that has existed for three decades—regulation will likely not come fast enough to guide even the later-stage promises of brain-machine interfaces. Establishing proper cultural norms at the outset and careful consideration of technologies is warranted.

FOREIGN COLLABORATION

As noted earlier, useful products emerge from neuroscience only after scientific issues have been resolved and engineering challenges have been met. Scientific research in neuroscience is in effect precompetitive, and this remains the major roadblock for most useful products. This point suggests that the primary capital in neuroscience is human expertise, and that future success continues to depend on the United States being the best place for international scientists to train, conduct research, and use their own expertise to train the next generation of scientists. Against this backdrop, the apparent targeting of US scientists with personal and familial links to China raises concerns,[28] and the United States only loses if these scientists leave and move their labs to China.

ETHICAL FRAMEWORKS

Neuroscience research naturally raises several ethical concerns that merit careful ongoing discussion and monitoring. Chief among these is human subjects research, of which there are many existing frameworks and regulations that guide neuroscience research 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. Society for Neuroscience, "Neuroscience Core Concepts," accessed August 15, 2023, https://www.brainfacts.org/core-concepts.

2. Rohit Bhisey, "Neuroscience Market Growing at a CAGR of 3.9% during the Forecast Period (2020–2030)," BioSpace, September 2, 2022, https://www.biospace.com/article/neuroscience-market-growing-at-a-cagr-of-3-9-percent-during-the-forecast-period-2020-2030-/; IMARC Group, "Neuroscience Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2023–2028," accessed August 15, 2023, https://www.imarcgroup.com/neuroscience-market.

3. 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, https://doi.org/10.1073/pnas.172399499.

4. Stanford Medicine, "The Stanford Artificial Retina Project," accessed August 30, 2023, https://med.stanford.edu/artificial-retina.html.

5. Jerry Tang et al., "Semantic Reconstruction of Continuous Language from Non-Invasive Brain Recordings," *Nature Neuroscience* 26 (May 2023): 858–66, https://doi.org/10.1038/s41593-023-01304-9; Sigal Samuel, "Mind-Reading Technology Has Arrived," Vox, May 4, 2023, https://www.vox.com/future-perfect/2023/5/4/23708162/neurotechnology-mind-reading-brain-neuralink-brain-computer-interface.

6. Philip A. Kragel and Kevin S. LaBar, "Decoding the Nature of Emotion in the Brain," *Trends in Cognitive Science* 20, no. 6 (June 2016): 444–45, https://doi.org/10.1016/j.tics.2016.03.011.

7. Winston Wong, "Economic Burden of Alzheimer Disease and Managed Care Considerations," *American Journal of Managed Care* 26, no. 8 (August 2020): S177–83, https://doi.org/10.37765/ajmc.2020.88482.

8. Alzheimer's Association, "2023 Alzheimer's Disease Facts and Figures," March 14, 2023, https://doi.org/10.1002/alz.13016; A. W. Willis et al., "Incidence of Parkinson's Disease in North America," *npj Parkinson's Disease* 8, no. 170 (December 2022), https://doi.org/10.1038/s41531-022-00410-y.

9. Willis et al., "Incidence of Parkinson's Disease."

10. Sara Reardon, "FDA Approves Alzheimer's Drug Lecanemab amid Safety Concerns," *Nature* 613 (January 2023): 227–28, https://doi.org/10.1038/d41586-023-00030-3.

11. Reardon, "FDA Approves Alzheimer's Drug Lecanemab."

12. Vassilis E. Koliatsos and Vani Rao, "The Behavioral Neuroscience of Traumatic Brain Injury," *Psychiatric Clinics of North America* 43, no. 2 (June 2020): 305–30, https://doi.org/10.1016/j.psc.2020.02.009.

13. Wu Tsai Neurosciences Institute, "NeuroChoice Initiative (Phase 2)," Stanford University, accessed August 30, 2023, https://neuroscience.stanford.edu/research/funded-research/neurochoice.

14. Low end: Pew Charitable Trust, "The High Price of the Opioid Crisis, 2021," 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 (April 16, 2021): 541–46, http://dx.doi.org/10.15585/mmwr.mm7015a1.

15. National Institute on Drug Abuse, "Drug Overdose Death Rates,", National Institutes of Health, June 30, 2023, https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates.

16. Overdose included in the accidental death statistic in Centers for Disease Control and Prevention, "Leading Causes of Death," last modified January 18, 2023, https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm.

17. US Food and Drug Administration, "FDA Takes Steps Aimed at Fostering Development of Non-Addictive Alternatives to Opioids for Acute Pain Management," February 9, 2022, https://www.fda.gov/news-events/press-announcements/fda-takes-steps-aimed-fostering-development-non-addictive-alternatives-opioids-acute-pain-management.

18. Haider J. Warraich, "What an 1890s Opioid Epidemic Can Teach Us About Ending Addiction Today," Stat, February 11, 2020, https://www.statnews.com/2020/02/11/1890s-opioid-epidemic-teach-us-about-addiction-today.

19. Matthew B. Pomrenze et al., "Modulation of 5-HT Release by Dynorphin Mediates Social Deficits during Opioid Withdrawal," *Neuron* 110, no. 24 (December 2022): 4125–43, https://doi.org/10.1016/j.neuron.2022.09.024; Gordy Slack, "Social Aversion during Opioid Withdrawal Reflects Block 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.

20. Emily Petrus, NINDS, and Laura Stephenson Carter, "Opioid Addiction and Chronic Pain: NIH Pain Consortium Symposium Highlights," *NIH Catalyst*, National Institutes of Health, April 6, 2022, https://irp.nih.gov/catalyst/26/5/opioid-addiction-and-chronic-pain.

21. S. Michaela Rikard et al., "Chronic Pain among Adults—United States, 2019–2021," *Morbidity and Mortality Weekly Report* 72, no. 15 (April 14, 2023): 379–85, https://dx.doi.org/10.15585/mmwr.mm7215a1.

22. M. M. Heinricher et al., "Descending Control of Nociception: Specificity, Recruitment and Plasticity," *Brain Research Reviews* 60, no. 1 (April 2009): 214–25, https://doi.org/10.1016/j.brainresrev.2008.12.009.

23. Lindsay M. Lueptow, Amanda K. Fakira, and Erin N. Bobeck, “The Contribution of the Descending Pain Modulatory Pathway in Opioid Tolerance,” *Frontiers in Neuroscience* 12, no. 886 (November 2018), https://doi.org/10.3389/fnins.2018.00886.

24. Seyyed Salman Alavi et al., “Behavioral Addiction versus Substance Addiction: Correspondence of Psychiatric and Psychological Views,” *International Journal of Preventive Medicine* 3, no. 4 (April 2012): 290–94, https://pubmed.ncbi.nlm.nih.gov/22624087.

25. Society for Neuroscience, “Attendance Statistics: Meeting Attendance,” accessed August 30, 2023, https://www.sfn.org/meetings/attendance-statistics.

26. Stefan Sykes, “U.S. Health Officials Want to Loosen Marijuana Restrictions: Here’s What It Means,” CNBC, August 31, 2023, https://www.cnbc.com/2023/08/31/hhs-wants-to-reclassify-marijuana-what-it-means.html.

27. Roper v. Simmons, 543 U.S. 551 (2005).

28. 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.

06

NUCLEAR TECHNOLOGIES

KEY TAKEAWAYS

- Nuclear fission offers a promising carbon-free power source that is already in use but faces safety and proliferation concerns, economic obstacles, and significant policy challenges to address long-term radioactive waste disposal.
- Nuclear fusion recently achieved an important milestone by demonstrating energy gain in the laboratory for the first time. However, further research breakthroughs must be achieved in the coming decades before fusion can be technically viable as an energy alternative.
- Many believe that small modular reactors (SMRs) are the most promising way to proceed with nuclear power, but some nuclear experts have noted that SMRs do not solve the radioactive waste disposal problem.

Nuclear technologies include those for nuclear energy production, nuclear medicine, and nuclear weapons. This chapter focuses on nuclear energy, which exploits the energy present in the nuclei of atoms. Fission and fusion are the two ways to tap that energy. Both fission and fusion reactions produce large amounts of heat, which can then be used to generate steam. Steam in large amounts can be used to drive turbines that produce electricity.

Overview: Nuclear Fission Technology

Nuclear fission is the process of striking the nucleus of a fissile isotope such as Uranium-235 with a neutron, causing it to split into smaller nuclei of lighter elements—and release energy. The split also releases neutrons that can go on to split other fissionable nuclei, resulting in a chain reaction. If the chain reaction is uncontrolled, what happens is a nuclear explosion. But a tightly controlled nuclear chain reaction can produce a continuous release of energy at low levels that can generate electricity.

Fission-driven power generation was first demonstrated in 1951.[1] Nuclear (fission) reactors can produce electric power in vast amounts without carbon emissions, but the reaction also produces radioactive by-products that must be safely managed for tens of thousands of years.

The spread of fission reactors can also raise concerns about the spread of nuclear weapons, since knowledge and infrastructure to design, build, and operate a nuclear power plant overlap substantially with what is needed to build nuclear weapons. In this view, research on new nuclear reactors whitewashes the nuclear power–nuclear weapons connection. Others believe that the proliferation risks can be minimized to the extent that fission reactors are a viable option for emissions-free energy.

Overall, in the last couple of years, the global capacity for nuclear reactors to generate electric power has declined slightly. The new nuclear reactors coming online, mostly in Asia, are unable to replace the capacity loss due to aging nuclear reactors being decommissioned in the West.

In addition, the United States does not offer competitive exports of nuclear power plants. Although there are some exports from the United Arab Emirates and South Korea, Russia dominates the global market for nuclear reactor exports. South Korea has a single design and more expertise in industrial manufacturing, allowing it to maintain low costs. Russia's state-owned Rosatom nuclear energy corporation has better financing and offers a more complete fuel provision and waste disposal.

Commercial reactors offer other potential applications as well, since two-thirds of the energy converted from nuclear reactors are released as heat to the environment. This energy could be harnessed to use for heat demands in other industrial processes, notably in desalination plants, metal refining, and hydrogen generation. These use cases are still in the process of development, with the Department of Energy (DOE) supporting US nuclear energy companies.

Commercial nuclear energy is used exclusively for electricity generation. In 2020, nuclear energy provided 10 percent of global electricity generation, making it the second-largest source of low-carbon electricity, behind hydroelectricity.[2] In the United States, nuclear power contributes 18.2 percent of electricity generation, the largest source of carbon-free electricity.[3]

Research and development in nuclear energy focuses on new reactor designs that may reduce nuclear fuel requirements, provide improved safety, and be less expensive to build and operate. R&D is also exploring approaches to disposal and long-term management of radioactive waste resulting from reactor operation.

Key Developments: Fission

New Reactor Designs

Advanced reactors could potentially offer a variety of benefits for:[4]

- **Safety** Advanced reactors could offer passive safety features that do not require direct human intervention to be activated or reactor operation at lower pressure that can reduce the risk of explosion.
- **Industrial decarbonization** Some advanced nuclear reactors can generate enough heat for

industrial processes that would otherwise be generated by fossil fuels.

- **Radioactive waste reduction** Some designs seek to reduce the amount of long-term radioactive waste produced in the power generation process; however, no reactor produces no radioactive waste at all.

One new reactor design gaining traction is the small modular reactor (SMR). These reactors generate less than 300 megawatts of electricity, about 25 to 30 percent the capacity of a conventional reactor. Smaller than conventional reactors, SMRs have the benefit that they can be mass produced in factories and transported to installation sites. Timelines for approval could be significantly reduced because the design of any given SMR would have to be reviewed only once. Multiple SMRs could support large power plants, while single SMRs could power smaller ones.[5]

On the other hand, SMRs are currently at the demonstration and licensing phase and hence remain an unproven technology. Moreover, while SMRs are designed to reduce capital costs, a large fraction of an SMR's cost goes toward preparing the site, which means that the use of an SMR saves 30 to 40 percent in cost—but produces 70 percent less power. SMRs also generate a greater volume of waste per unit of energy produced as compared to larger reactors.[6]

Fuel for New Reactors

Uranium ore consists of about 99.3 percent Uranium-238 and 0.7 percent Uranium-235. For use in today's commercial light-water reactors, uranium must be "enriched" to increase the concentration of U-235 from 0.7 percent to about 3 to 5 percent, making it "low-enriched" uranium. Most new reactor designs, however, call for the use of uranium fuel enriched with U-235 at a level between 5 percent and 20 percent, fuel known as high-assay low-enriched uranium (HALEU).[7] However, HALEU is unavailable at a commercial scale, and projections suggest that more than 40 metric tons of HALEU will be needed before the end of the decade in these advanced reactors should they actually be deployed.[8] US government–supported research is underway to develop processes to produce commercially viable HALEU. These processes use spent nuclear fuel from government-owned research reactors to produce small amounts of HALEU—the first steps in the creation of a domestic HALEU supply for advanced nuclear reactors.

More than 90 percent of the uranium used in US nuclear reactors is imported; Kazakhstan and Russia account for nearly half of all US uranium consumption, while Canada and Australia account for about 30 percent.[9] One approach to eliminate the need for uranium imports is to extract uranium from seawater. In total, seawater contains hundreds of times more uranium than is on land, but extracting it for use in nuclear power generation is challenging due to its low concentration and the high-salinity background.[10] As noted by Stanford professor Steven Chu, former US secretary of energy under President Barack Obama: "Seawater extraction gives countries that don't have land-based uranium the security that comes from knowing they'll have the raw material to meet their energy needs."[11]

Nuclear Waste Disposal

Radioactive nuclear waste can be differentiated between high-level and low-level waste based on how long it takes before the waste decays and is no longer hazardous. High-level waste includes "spent," or used, nuclear fuel and waste generated from the reprocessing of spent fuel. Low-level waste includes items that have come in contact with radioactive materials; such items include paper, rags, plastic bags, or clothing. In terms of overall volume, less than 1 percent of existing radioactive waste is high level; about 4 percent is intermediate level; and around 95 percent is low level. This low-level waste can take a few years or decades to decay, while high-level waste can take upward of tens of thousands of years.

Managing nuclear waste requires answering two primary questions: how to store it and where to store it. Low-level nuclear waste is most often stored in metal drums; high-level waste is by law turned into glass, or vitrified, to immobilize it and then stored in containers. But by far the most controversial issue in waste management is where to store it.

After cooling for years in water, low-level waste is moved to dry storage aboveground. High-level waste requires deep underground repositories to isolate it for thousands of years. However, identifying suitable sites is highly contested, despite a broad consensus that such waste should be stored underground (as opposed to burying it at sea, for example). Because it must be stored for so long, a geologically stable environment is needed to ensure that earth movements do not disturb the waste repositories, and a dry environment is needed to ensure that running water does not leach away waste and transport it from the disposal site. This is a possibility because long-lived fission products and some activation products have geochemical properties that prevent them from binding onto the surfaces of minerals that would otherwise immobilize them in place.

Finally, the idea of transmuting the radioactive elements in nuclear waste into less dangerous elements is occasionally floated. Natural transmutation for nuclear waste materials occurs over time but takes hundreds of thousands or millions of years. Speeding up this process entails subjecting the nuclear waste elements to some other nuclear process to effect a transformation and has been demonstrated on the atomic scale in the laboratory—but never on a scale necessary to deal with the 86,000 tons of high-level radioactive waste now being stored temporarily in aboveground sites.

Over the Horizon: Fission

Impact of New Reactor Designs

Generation IV nuclear reactors are proposed reactors that are more advanced than the Generation III and III+ reactors in use today. Generation IV reactors seek to improve sustainability, economics, safety and reliability, proliferation resistance, and physical protection. Some of the technical goals of such reactors include increased efficiency of electricity generation; generation and capture of process heat to be used in other thermal applications, such as the production of hydrogen; increased safety; and reduced production of waste materials.

Generation IV reactors are characterized by their coolants, which can be water, helium, liquid metal, or molten salt, and by whether they operate with moderated (slower) or unmoderated (faster) neutrons. Reactors using moderated (or thermal) neutrons can operate with low-enriched uranium fuel, which presents a lower risk of nuclear weapons proliferation. Reactors using unmoderated (or fast) neutrons must use HALEU but are able to generate more power per unit of fuel.

According to the US National Academies of Sciences, Engineering, and Medicine, "advanced nuclear technologies likely will not be able to markedly contribute to electricity generation until the 2030s at the earliest."[12] Nevertheless, they may compete with other energy technologies in the long term.

Challenges of Innovation and Implementation

Bridging the gap between innovation and implementation remains a challenge for advanced Generation IV reactors. The design for such reactors has been on the books for many years, and the scientific theory of nuclear power generation and the engineering know-how to build nuclear plants have also been available for many years. Nevertheless, concerns over matters such as cost and safety have largely prevented any action being taken toward deployment. China connected the first Generation IV reactor—a demonstration project—to its power grid on December 20, 2021,[13] but no other Generation IV plants are known to be under construction anywhere else in the world.

Policy, Legal, and Regulatory issues

Waste management There is no enduring US plan for a long-term "permanent" solution to disposing of nuclear waste, with essentially all civilian nuclear waste being "temporarily" stored on-site at nuclear power plants. The one site that was seriously proposed for permanent storage at Yucca Mountain was shut down in 2010. The Obama administration cited opposition from the State of Nevada in suspending the Yucca Mountain Project. There are no new fuel disposal or storage facilities for long-term US use currently in development by the DOE, although at this writing, Finland is expected to open its Onkalo site for permanent storage of spent fuel in 2024.[14] Two private-sector facilities for interim storage (Consolidated Interim Storage Facilities) have been proposed in Texas and New Mexico, but host states have opposed the Nuclear Regulatory Commission's licensing of these facilities.[15] Both states have received NRC licensing, but the approval of the Texas site was vacated by the US Court of Appeals for the Fifth Circuit.[16] The amount of US high-level nuclear waste to be managed is today around 86,000 tons and grows at the rate of an additional 2,000 tons per year—which makes management of such waste an important public policy concern.

Economics Nuclear energy and economics are intrinsically linked, with both capital costs and the operating costs of energy production directly influencing the economy's health and competitiveness. At the construction phase, conventional nuclear power plants have experienced significant construction cost overruns. The construction of new fission power plants faces delays due to Nuclear Regulatory Commission intervention during construction, state rules that delay permitting, and a lack of advocates for new nuclear plants. At the operational phase, nuclear-generated electricity is not cost-competitive due to high operating costs. In the United States, the cost of upgrades for older nuclear reactors and the relative marginal cost of nuclear compared to wind and solar (nuclear has higher marginal cost) have made nuclear power plants less economically feasible than other sources of renewable energy.

Timescale Recognizing the urgency of reducing greenhouse gas emissions and the time it takes to approve and build new reactor designs safely, it is unclear whether a sufficient number of nuclear reactors can become operational in time. According to the International Energy Agency, 439 nuclear power reactors were in operation in 2021, with a combined capacity of 413 gigawatts, which avoids 1.5 gigatons of global emissions per year.[17] Considering that global emissions in 2022 reached 36.8 gigatons,[18] doubling the number of reactors would only reduce global emissions by 4 percent (assuming efficiency remains the same). The median construction time of nuclear reactors connected to the grid in 2021 was eighty-eight months.[19] In the United States, the various approval processes take about sixty months.[20]

There is no enduring US plan for a long-term "permanent" solution to disposing of nuclear waste.

All in all, a twelve-year period from initiation of the approval process to grid connection does not seem excessive.

Fuel supply For fission in new nuclear reactors, the only commercial source of HALEU today is Russia, and the security and reliability of Russia as a source is not assured. Although the US government is undertaking research that might result in the availability of a domestic supply, environmental and other land-use issues might inhibit the development and deployment of facilities to produce HALEU.

Overview: Nuclear Fusion Technology

Fusion is another physical process that produces massive amounts of energy from atoms. Instead of splitting atoms like fission, fusion occurs when two atomic nuclei collide together to form a heavier nucleus. Substantial amounts of energy, several times greater than fission, are released without any long-lived radioactive waste. Fusion is the source of energy in a thermonuclear bomb—and the sun. As with nuclear fission, the hope is that fusion can be controlled to drive electrical generators.

Fusion energy comes from the fusion of deuterium (D) and tritium (T), both isotopes of hydrogen. Deuterium is common in seawater, but tritium is radioactive and, because of its short half-life of twelve years, is not found in nature and thus must be manufactured. The D-T reaction produces a helium-4 nucleus and a fast neutron.

Fusion energy is still in the R&D stage. There are two approaches in serious fusion research today, and both attempt to solve what is known as the confinement problem.[21]

The confinement problem refers to the challenge of keeping a fusion fuel—typically a mix of hydrogen isotopes like deuterium and tritium—at the necessary high temperatures and pressures long enough for a significant number of nuclear fusion reactions to occur. Because fusion involves "fusing" two nuclei together, the fusion reaction must overcome the repulsive forces between two charged nuclei—and the only known way to do that is to have the nuclei moving at very high speeds, corresponding to being at a very high temperature.

One way to confine the fuel is to use powerful magnets to trap a high-temperature plasma of deuterium and tritium, a process known as magnetic confinement fusion (MCF). These magnets keep the hot plasma away from the containment vessel walls, aiming to maintain the necessary high temperatures and densities for the fusion reactions to occur in sufficient frequency. The engineering challenge is to ensure stability of the plasma and maintain confinement conditions sufficiently long enough (several seconds) for a net positive energy output, as plasma instabilities can disrupt this process.

A second way—inertial confinement fusion (ICF)—calls for the very rapid compression of a fuel pellet using lasers or ion beams, causing the fuel pellet to implode. The beams hit the pellet's surface simultaneously, causing the pellet's outer layer to explode, thus driving the rest of the pellet toward its center. The beams are very powerful but illuminate the pellet for a short time, around 20 nanoseconds, during which the pellet is compressed. When an adequate degree of compression has been achieved, ignition of the fuel begins, and for an even shorter time of about 100 picoseconds, the compressed fuel—now a very hot plasma—does not have a chance to move very much because of its own inertia—hence the name inertial confinement. Here, the engineering challenge is ensuring that the beams hit the pellet simultaneously in a symmetrical manner, and the rate at which pellets can be dropped and imploded determines the rate at which energy is released.

In a typical conceptual design for a fusion reactor, a pellet might be dropped ten to twenty times per second, requiring the illuminating lasers to fire that often. The laser energy incident on the pellet would be a couple of megajoules, and the fusion reaction would produce around 100 to 150 megajoules, for an energy gain of fifty to one hundred times. (Energy gain is the ratio of the energy produced to the energy used to initiate the fusion reaction. It is important because only if the gain is greater than one is the reaction producing net energy.) Important engineering challenges include building facilities for mass production of fusion pellets, as a single reactor might use a million pellets per day. Other challenges include reducing the cost of pellets (a target goal might be 10 to 50 cents per pellet), developing lasers that can fire ten to twenty times per second, and finding structural materials for building the reactor that can acceptably withstand the fast neutrons that are emitted in the fusion reaction.

For fusion to be a viable energy source, the confinement strategy must allow more energy to be produced from the fusion reactions than the energy invested in initiating those reactions (i.e., the energy gain must be greater than one). Achieving a net positive energy output while managing the confinement challenges is a central problem in fusion research.

Research on nuclear fusion is performed in a number of government, commercial, and academic institutions. Government involvement occurs in a number of national laboratories. A few dozen private-sector companies are active in fusion and a dozen or so universities are also involved. Funding for fusion research comes from private capital and US government coffers: research for fusion for energy production received $763 million from the US government for fiscal year 2023 and fusion research related to nuclear weapons received an additional $630 million.[22] Private companies declared funding of $4.7 billion in the 2022 calendar year.[23]

Key Developments: Fusion

A milestone was reached in December 2022, when the National Ignition Facility (NIF) at Lawrence Livermore National Lab reached better than "breakeven" in an ICF experiment—in other words, the point at which more energy was released by a nuclear fusion reaction than the proximate energy used to initiate the reaction;[24] in this experiment, the energy released was 1.53 times the proximate energy (i.e., the energy gain was 1.53). A second demonstration of "better-than-breakeven" was repeated in July 2023.[25]

In both cases, the proximate energy was the energy used by the lasers involved in initiation. However, these experiments did not come anywhere near to breakeven if the energy inputs to the lasers are considered. Nevertheless, these experiments have spurred interest in the field of nuclear fusion, especially among the large number of start-ups in this arena. While most of the investment in such companies comes from venture capital firms, the Department of Energy has outlined plans for public-private partnerships to develop on-grid fusion energy within the next few decades.

Over the Horizon: Fusion

Impact of New Technology

The fusion energy future faces many technical research challenges, including:

- **Reactor configuration** The feasibility of fusion as a power source depends on solving the confinement problem, and we don't know if magnetic confinement or inertial confinement will prove to be feasible methods in the long run.

- **Availability of tritium** Because tritium is not found in appreciable amounts in nature, it must be manufactured. Tritium can be produced in fission nuclear reactors by subjecting lithium to neutron irradiation in the reactor's core. The United States has not needed to produce tritium for several decades, but if fusion power becomes commercially viable, it will have to obtain sufficient supplies.

- **Fabrication of fuel** Fusion reactors require the preparation of the D-T mixture into geometric forms that easily absorb the energy needed to push the nuclei together.

- **Materials** The physical structures housing fusion reactions are subject to damaging bombardment, since most proposed fusion reactions, including deuterium-tritium, release high-energy neutrons. New materials are needed that are more neutron resistant.

Challenges of Innovation and Implementation

Some press accounts of the genuine breakthrough in achieving scientific breakeven gave the impression that the experiment suggested that practical fusion energy was "just around the corner." Even the most optimistic private investors in fusion do not believe it is any closer than ten to fifteen years away.

Policy, Legal, and Regulatory Issues

Nuclear proliferation Fusion power plants will generate fast neutrons in addition to producing useful heat. These neutrons can be used in principle to transmute certain elements into material that can be used to make fission weapons. One study on this topic acknowledges some proliferation risk but concludes that "proliferation risk from fusion systems can be much lower than the equivalent risk from fission systems, provided commercial fusion systems are designed to accommodate appropriate safeguards."[26]

Waste management The primary waste products from nuclear fusion are the materials irradiated by the intense neutron radiation produced in the fusion reaction. The neutrons serve to transmute the elements in the original materials into other elements, and often these "activation products" are radioactive. However, they generally do not remain dangerous for nearly as long as the waste products from fission reactors.

NOTES

1. Office of Nuclear Energy, "9 Notable Facts About the World's First Nuclear Power Plant—EBR-I," US Department of Energy, June 18, 2019, https://www.energy.gov/ne/articles/9-notable-facts-about-worlds-first-nuclear-power-plant-ebr-i.

2. International Energy Agency, "Nuclear Power," last modified July 11, 2023, https://www.iea.org/reports/nuclear-electricity.

3. US Energy Information Administration, "Frequently Asked Questions: What Is U.S. Electricity Generation by Energy Source?," last modified March 2, 2023, https://www.eia.gov/tools/faqs/faq.php?id=427&t=3.

4. Vincent Gonzales and Lauren Dunlap, "Advanced Nuclear Reactors 101," *Resources for the Future*, March 26, 2021, https://www.rff.org/publications/explainers/advanced-nuclear-reactors-101.

5. US Energy Information Administration, "Nuclear Explained: Nuclear Power Plants," last modified August 7, 2023, https://www.eia.gov/energyexplained/nuclear/nuclear-power-plants-types-of-reactors.php.

6. Lindsay M. Krall, Allison M. Macfarlane, and Rodney C. Ewing, "Nuclear Waste from Small Modular Reactors," *Proceedings of the National Academy of Sciences* 119, no. 23 (May 2022), https://doi.org/10.1073.pnas.2111833119. According to this study, the volume of waste per unit of energy produced is anywhere from two to thirty times that from a typical large reactor.

7. Office of Nuclear Energy, "What Is High-Assay Low-Enriched Uranium (HALEU)?," US Department of Energy, April 7, 2020, https://www.energy.gov/ne/articles/what-high-assay-low-enriched-uranium-haleu.

8. US Department of Energy, "DOE Announces Cost-Shared Award for First-Ever Domestic Production of HALEU for Advanced Nuclear Reactors," November 10, 2022, https://www.energy.gov/articles/doe-announces-cost-shared-award-first-ever-domestic-production-haleu-advanced-nuclear.

9. US Energy Information Administration, "Nuclear Explained: Where Our Uranium Comes From," last modified August 23, 2023, https://www.eia.gov/energyexplained/nuclear/where-our-uranium-comes-from.php.

10. Chong Liu et al., "A Half-Wave Rectified Alternating Current Electrochemical Method for Uranium Extraction from Seawater," *Nature Energy* 2, no. 17007 (February 2017), https://doi.org/10.1038/nenergy.2017.7.

11. Tom Abate, "How to Extract Uranium from Seawater for Nuclear Power," Stanford Engineering, Stanford University, February 17, 2017, https://engineering.stanford.edu/magazine/article/how-extract-uranium-seawater-nuclear-power.

12. Joshua Blatt, "U.S. Should Begin Laying the Foundation for New and Advanced Nuclear Reactors, Says New Report," National Academies of Science Engineering, and Medicine, April 27, 2023, https://www.nationalacademies.org/news/2023/04/u-s-should-begin-laying-the-foundation-for-new-and-advanced-nuclear-reactors-says-new-report.

13. Sonal Patel, "China Starts Up First Fourth-Generation Nuclear Reactor," *Power*, February 1, 2022, https://www.powermag.com/china-starts-up-first-fourth-generation-nuclear-reactor.

14. YLE News, "Finland Soon Home to World's First Permanent Nuclear Waste Site," January 16, 2023, https://yle.fi/a/74-20013058.

15. Congressional Research Service, *Civilian Nuclear Waste Disposal*, September 17, 2021, https://crsreports.congress.gov/product/pdf/RL/RL33461.

16. State of Texas v. Nuclear Regulatory Commission, 21-60743 (5th Cir. 2023).

17. Peter Fraser et al., *Nuclear Power and Secure Energy Transitions: From Today's Challenges to Tomorrow's Clean Energy Systems*, International Energy Agency, September 2022, 7, 14, https://iea.blob.core.windows.net/assets/016228e1-42bd-4ca7-bad9-a227c4a40b04/NuclearPowerandSecureEnergyTransitions.pdf.

18. International Energy Agency, *CO_2 Emissions in 2022—Analysis*, March 2023, https://www.iea.org/reports/co2-emissions-in-2022.

19. Statista, "Median Construction Time Required for Nuclear Reactors Worldwide from 1981 to 2022 (in months)," accessed August 30, 2023, https://www-statista-com/statistics/712841/median-construction-time-for-reactors-since-1981.

20. US Energy Information Administration, "Nuclear Explained: U.S. Nuclear Industry," last modified August 24, 2023, https://www.eia.gov/energyexplained/nuclear/us-nuclear-industry.php.

21. A third approach, cold fusion, was announced in 1989. Two chemists claimed to have observed a fusion reaction that could take place at room temperature. However, these observations have not been repeated, and most scientists today believe that such reactions are not possible and that the results observed in 1989 reflected error rather than breakthrough.

22. Fusion Industry Association, "Congress Provides Record Funding for Fusion Energy," accessed August 30, 2023, https://www.fusionindustryassociation.org/congress-provides-record-funding-for-fusion-energy.

23. David Dalton, "Private Funding Has Increased by $2.8 Billion in a Year, Says Industry Survey," NucNet, July 15, 2022, https://www.nucnet.org/news/private-funding-has-increased-by-usd2-8billion-in-a-year-says-industry-survey-7-5-2022.

24. Breanna Bishop, "Lawrence Livermore National Laboratory Achieves Fusion Ignition," Lawrence Livermore National Laboratory, December 14, 2022, https://www.llnl.gov/news/lawrence-livermore-national-laboratory-achieves-fusion-ignition.

25. Ben Brasch, Kyle Rempfer, and Shannon Osaka, "U.S. Lab Says It Repeated Fusion Energy Feat—with Higher Yield," *Washington Post*, August 7, 2023, https://www.washingtonpost.com/climate-solutions/2023/08/06/nuclear-fusion-net-energy-gain-higher-yield.

26. R. J. Goldston, A. Glaser, and A. F. Ross, *Proliferation Risks of Fusion Energy: Clandestine Production, Covert Production, and Breakout*, Office of Scientific and Technical Information, US Department of Energy, August 13, 2009, https://www.osti.gov/biblio/962921.

07

ROBOTICS

KEY TAKEAWAYS

- Although robots today are mostly used for the Three Ds (dull, dirty, or dangerous tasks), in the future they could be used for almost any task involving physical presence, because of recent advances in AI, decreasing costs of mobile component technologies (e.g., cameras in smartphones), and designs enabled by new materials and structures.
- Robotics has and will transform many industries through elimination, modification, or creation of jobs and functions.
- Understanding and communicating how robots will affect people's lives directly in their physical spaces (e.g., security robots in malls) as well as more existentially (e.g., transitioning jobs like truck driving from human-driven to autonomous vehicles) will shape how the United States accepts and benefits from robotic technologies.

Overview

What is a robot? Researchers do not universally agree on the definition of a robot, but a consensus seems to have emerged that at the very least a robot is a human-made physical entity with ways of sensing itself or the world around it and ways of creating physical effects on that world.[1]

The global robotics market is estimated at $25 billion today and poised for strong growth over the next decade. While projections vary greatly, some consultancies estimate that the global market could be worth between $160 billion and $260 billion by 2030.[2] The adoption of professional service robots is expected to drive this growth. These types of robots currently occupy a small sliver of the market, but as technology continues to improve, they will find greater adoption in industries like medicine, agriculture, and construction. The United States is projected to maintain a plurality of global robotics revenues for the next few years.

Importantly, robots must integrate many different component technologies to combine perception of the environment with action on the environment. Perception requires generating a representation of the robot's environment and interaction with it. 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 (e.g., enabling visual or other perception) and then integrating them within a physical or mechanical structure to perform the robot's intended tasks (e.g., using perception to guide motions and actions) in different settings in a given environment. The physical structure could be regarded as a robot's body. Further, different types of robots operate in different environments (e.g., factories, homes, and even space), and each environment raises distinct complexities beyond just technical performance. For example, working alongside humans raises critical issues of safety and liability.

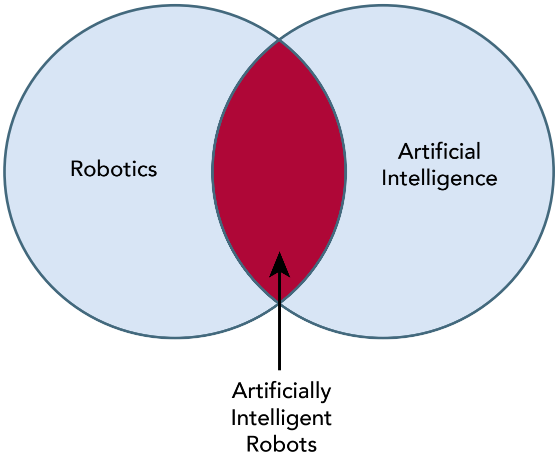
The dependence of robotics on many different component technologies and nontechnological considerations has an important practical consequence—it takes a huge interdisciplinary effort, not just from technologists but also from experts in other fields, to move from a working prototype in a research lab to a useful functional robot in the market.

Important component technologies include:

- **Actuators** These components enable movement (e.g., motors, grasping appendages). Today, mechanical actuators are typically rigid, restricting the environments where robots can operate. These inflexible actuators make operating in confined spaces and on irregular terrain or performing dexterous movements in unknown environments challenging.
- **Sensors** These receive real-time input about the immediate physical environment of the robot and the robot's own configuration. Such inputs inform decisions about what the robot should be doing in the next moment in time.
- **Control systems** These components decide how actuators should move based on readings from sensors.
- **Materials** Constructed of rigid materials and with joints based on ordinary bearings, traditional robots interact with their operating environments in highly prescribed and structured ways. "Soft" robots that are flexible and conform to the environment can offer better performance in the more unstructured and chaotic environments that characterize most of the world, but the construction of soft robots often entails the creation of new materials or structures.
- **Power sources** Tethered robots can be energized from a power source on the "mother ship" indefinitely, while untethered robots need self-contained power sources or sources that harvest energy from the environment. A common portable power source is batteries, which drain themselves quickly—too quickly for many practical applications of robots.
- **Real-time programming** As physical devices, robots operate in real time and their components must operate within the boundaries of physical timelines determined by the operation of the robot. An actuator that moves a tenth of a second too early may cause a robot hand to fail at grasping an object. Deviations in timing may have nothing to do with the programming of the real-time microprocessor, but rather occur because another subsystem in the robot failed to operate on time or because something unexpected happened in the robot's environment.

While some robots use computer vision and other types of AI for understanding their environments and decision making, robotics and AI do not always

FIGURE 7.1 Not all robots use artificial intelligence



go together (see figure 7.1). Some robots, for example, are tele-operated, requiring a human operator to control or direct most aspects of the robot's behavior.[3] Other robots can do a few things autonomously without AI, such as maintaining themselves in a fixed position or entering a "cruise control" mode, like the early versions of the Roomba, a series of robotic vacuum cleaners introduced in 2002. There are robots that rely on AI for extensive capabilities for autonomous behavior and decision making, such as the self-driving taxis approved for business in San Francisco in 2023. Robots with varying degrees of autonomy have been used in everything from delicate surgical procedures to space exploration.

Robots are used primarily for the Three Ds: human work that is dull, dirty, or dangerous. Human attention to a task waxes and wanes, whereas robots do not get bored. Human life and well-being are precious compared to the physical damage that a robot might experience in doing a task. And robots can survive in much more hazardous or extreme environments—cleaning up nuclear reactors and exploring Mars, for example—than humans can.[4]

Key Developments

Robots are used across a wide range of sectors in a variety of ways. Prominent examples include:

- **Manufacturing** Many assembly lines use stationary robots (see figure 7.2) to undertake repetitive tasks such as welding at high speeds and with great precision.[5] The environment around such robots is highly structured and carefully controlled to minimize the need for robot cognizance of surroundings. In most cases, these robots work in isolated cells without any physical interaction with humans.

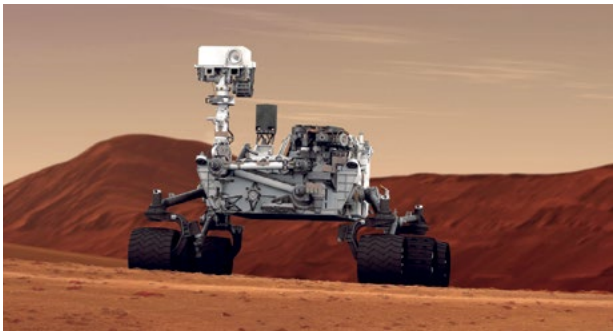
- **Warehouse logistics** Autonomous robots bring merchandise stored in very large warehouses to a central point for packaging.

- **Surgery** Mostly tele-operated today, surgical robots assist with minimally invasive surgery. Surgeons can reduce the size of the incisions that are needed for treatments and thus reduce surgical risks.[6] A surgical robot typically includes a camera arm and surgical instruments attached to mechanical arms controlled by a surgeon at a console operating the robot.

FIGURE 7.2 Automotive assembly line robots

Source: Carol M. Highsmith Collection, Library of Congress

FIGURE 7.3 Mars *Curiosity* rover



Source: NASA

FIGURE 7.4 *Sentry* autonomous underwater vehicle, Woods Hole Oceanographic Institution

Source: NOAA Ocean Exploration

FIGURE 7.5 Search-and-rescue robot

Source: National Institute for Standards and Technology

- **Science and exploration** Robots have been used to explore other planets,[7] the vast littoral ocean zones,[8] buildings,[9] the insect world,[10] and the human body.[11] Planetary rovers (see figure 7.3) and remotely operated underwater vehicles (see figure 7.4) are two examples of such robots.

- **Food production** Agricultural robots[12] can help harvest crops by picking fruit and maintain farmland by weeding. Drones can inexpensively survey farmland, and robot-operated greenhouses enable food production.

- **Disaster assistance robots** These robots can maneuver in collapsed spaces to reach victims underneath rubble, bringing communications, oxygen, food, or medicine (see figure 7.5).[13]

- **Security** Mobile robots in parking lots and buildings such as shopping malls provide telepresence for centrally located security personnel.

- **Military services** Robots have been developed that help to perform a variety of military services, including load transport, surveillance and reconnaissance, mine clearance, and armed sentry duty.

- **Transportation** Autonomous vehicles and trucks are the most common examples.[14]

Robots excel at working in structured environments like manufacturing plants where conditions are predictable. Their precision, speed, and tirelessness allow robots to surpass human performance in repetitive tasks on the assembly line. However, humans still outperform robots in unstructured or poorly structured environments. Constraints on power, fuel supplies (e.g., battery life), and sensors also limit the ability of robots to compete with humans.

Most important, human intelligence and adaptability give people a major advantage in chaotic, real-world environments. For example, robots struggle to navigate cluttered spaces or manipulate unfamiliar objects. Even advanced robots cannot yet match humans' intuitive understanding of physics and the ability to improvise solutions. Robots rely heavily on precise planning and control, so encountering small, unexpected changes may cause catastrophic failure. Humans, on the other hand, can apply their past experiences and reasoning to adapt to unfamiliar situations on the fly.

Over the Horizon

New Robotic Technologies

A growing direction in robotics is the one where robots and humans work together to capitalize on the advantages of each.

Advances in artificial intelligence and soft robotics may help robots become more capable in unstructured spaces. Machine-learning algorithms could enable robots to perform well in environments that they have never seen before or never been programmed to encounter. Flexible, nonrigid robots can perform in a variety of environments, even those inaccessible by humans.

Advances in robotics will be linked to advances in artificial intelligence, the decreasing cost of mobile components, and novel designs enabled by new materials. Researchers in robotics today[15] are working in areas such as:

Haptic technology for robots Sensations of touch and feel are useful in manipulation and many social interactions—for example, touch feedback is vital to a surgeon when palpating tissue. Haptic technology enables doctors to use touch as an input device to robotic surgical systems and tele-operated robots, facilitating more intuitive control and stronger physical connections with remote environments.[16]

Robotic movement through self-deformation Soft robots are safer for humans—a human–soft robot collision creates an impact between soft human tissue and a soft robot body and is less likely to result in injury. Soft robots' ability to deform also affords novel methods of locomotion and manipulation. Soft robot technology can be applied in haptic interfaces, search and rescue, and medicine.[17] Soft robots can also use materials that are "continuum, configurable, and adaptable with functionalities relying on high degrees of freedom shape morphing."[18]

Robot design New types of robots may be needed for operation in difficult-to-access or uncertain environments. For example, robots for space exploration include small free-flying robots to operate with space stations, robot manipulator systems for use on craft that orbit planets, rovers that can jump in low-gravity environments, and robots that can move by stepping with adhesive pads.[19]

Wearable robots A wearable robot such as an exoskeleton enhances human mobility during activities like walking and running and may assist people with physical impairments.[20] They can also be used to augment human muscle power (see figure 7.6), thereby enabling people to lift large loads.

Robotic manipulation Situations in which grasping an object is necessary often do not provide perfect

FIGURE 7.6 Sarcos Guardian XO Max wearable exoskeleton

Source: Sarcos Technology and Robotics Corporation

information about the object, the grasping device, and the relative positions of everything in the environment. So it's important to develop manipulation sequences that work with uncertainty.[21]

Bio-inspired robots These are designed based on fundamental biological principles and could include living components.[22]

Robot swarms These small, modular robotic components act in coordination as a team to perform tasks.[23]

Artificial intelligence for robotics This refers to the challenge of giving robots the ability to learn how to learn and exhibit commonsense, intelligent behaviors.

Human-robot interaction (HRI) HRI focuses on understanding, designing, and evaluating robots for use by or with humans.[24] Humans and robots working closely together, for example, may have unpredicted, unintentional physical contact that may disrupt robot operation and cause safety issues for the human.[25] Another challenge is the design of these interactions in ways that accommodate social norms, are natural and seamless, and allow robots to exhibit more familiar and comfortable behaviors to humans.

Challenges of Innovation and Implementation

Supply chain issues are one of the most important near-term infrastructure challenges in robotics. The robotics field involves the integration of multiple foundational technologies, which means progress is heavily reliant on global supply chains for parts such as chips and materials. The more far-flung and complex the international innovation supply chain, the more slowly innovation will move when disruptions to the supply chain occur.

To illustrate, DJI is a Chinese company that controls a large share of the airborne drone market. One important reason for this dominance is the entire supply chain for DJI drones is self-contained

Advances in robotics will be linked to advances in artificial intelligence, the decreasing cost of mobile components, and novel designs enabled by new materials.

within one region of China. Start-up companies in the United States working in this space are generally forced to turn to Chinese suppliers as the US supply chain for drones is fragmented, making it cumbersome, expensive, and slow to deliver.

Policy, Legal, and Regulatory Issues

ROBOTS AND THE FUTURE OF WORK

Robots have enormous potential to affect jobs involving physical labor and presence in much the same way that AI might affect jobs involving knowledge and expertise. When realized in the marketplace, robotics is likely to eliminate some job types, create new job types, and modify the responsibilities and duties of jobs that remain.

Consider truck driving, one of the few well-paying jobs available to Americans without a college degree. The profession is likely to experience the following:

Job elimination Long-haul truck driving is likely to be one of the first jobs eliminated when autonomous land vehicles become feasible because fuel costs can be significantly reduced when trucks drive close to one another on highways—possible only with machine-speed reaction times rather than human reaction times.

Job modification Many truck-driving jobs entail both long-haul and local driving. Automation may take on more responsibility for long-haul driving, but navigating local driving conditions in a far less structured environment will take much longer to automate, which means truck-driving jobs may become more local in nature. But whether more local truck-driving jobs will continue to offer comparable compensation to those jobs of today is not at all clear.

Job creation Robot repair and technician jobs are likely to be a major job category as robots gain traction in the economy. If self-driving trucks come to dominate the landscape, we might see positions in logistics and fleet management increase, although these kinds of jobs entail a different skill set than that required for driving trucks.

ACCOUNTABILITY, REGULATION, AND LIABILITY

Societies routinely hold people accountable for harmful actions. People who cause accidents that harm people or property face liability for that harm. Soldiers in war who kill civilians because they use their weapons indiscriminately are guilty of war crimes. But as robots assume roles that call for similar decision making, how should concepts of individual accountability evolve?

Some questions include:

- What parties should be held accountable for harm occurring when robots are involved and how should those determinations be made?
- How and to what extent, if any, do robots involved in incidents that harm people or destroy property disproportionately or improperly attract liability lawsuits?
- How can existing regulatory regimes for transportation safety and medical safety, for example, keep pace with evolving robotics technology?
- How and to what extent, if any, are lethal autonomous weapons ethically and morally permissible?
- How is the safety of robot operation best assured? How, if at all, should safety trade off against other performance objectives?
- What are the appropriate standards of performance for robots? Robots often perform tasks that humans also perform. Should the standard be that the robot does the task nearly perfectly? Or is it adequate to perform it better than a human being?

If the latter, should the reference human being be an average person or a person who performs the task much better than most other humans?

SOCIETAL ACCEPTANCE

For robots to be widely used throughout society, their presence and operations should not cause human discomfort, unease, or fear. Experience suggests that many humans are disturbed by robots that look like humans (see figure 7.7), such as in healthcare settings.

Some advocate the use of robots for eldercare, suggesting that an aging population will create demands for services that cannot be met by future pools of workers interested in and qualified to perform those jobs.[26] It has been suggested that robots could help the elderly care for themselves by providing emotional support or cognitive therapy; enabling remote access for doctors and nurses; and entertaining home dwellers, monitoring them for falls, and helping them with housekeeping, lifting, and bathing needs. By assuming part of the eldercare labor force, robots could allow a limited number of care workers to do their jobs more efficiently and easily.

However, in one study of robots for eldercare in a nursing home, several challenges emerged, including the imposition of additional work burdens on human caregivers, the need for close monitoring of the robots, and the displacement of social and communication-oriented tasks that reduced opportunities for human connection.[27]

FIGURE 7.7 A very human-looking robot

Source: International Telecommunication Union, via Wikimedia Commons, CC BY 2.0

ROBOETHICS

Roboethics addresses what ethical guidance should be programmed into robots so they do not behave in ways that humans regard as unethical or lead to unethical outcomes. A related definition suggests that roboethics refers not to the ethics of robots but rather to the human ethics of robot designers, manufacturers, and users.[28]

One particularly clear example of roboethics arises in a military context of lethal autonomous weapons systems that can select a target and then act to destroy it without human intervention. Given military pressures to act more rapidly on the battlefield, many observers believe that such decisions will inevitably be turned over to computers. Others recoil at the prospect, arguing it is ethically wrong to make decisions in matters of life and death without human input.

But ethical issues also arise in more benign contexts. For example, what are the ethics of large-scale deployments of eldercare robots? Granting that many of the issues faced by an aging population relate to physical needs that either robots or home health caregivers could meet, what of their emotional needs? How and to what extent, if any, can interactions with an artificially intelligent eldercare robot provide comfort, empathy, and compassion for the person under its care?

PRIVACY AND SECURITY

Robots—especially mobile robots—often raise privacy and security concerns in unexpected ways. Cameras mounted on drones appear to raise many more privacy concerns than stationary cameras. Drones have also recently demonstrated the ability to open closed doors and therefore fly into areas that were once guaranteed to be drone-free.

IMPORTANCE OF MULTIDISCIPLINARY CENTERS

Robotics is a multidisciplinary field. Key developments will occur not only in labs that are studying "robotics" but also in those studying the underlying technologies that make robotics possible. These key fields include materials science, computer science, and artificial intelligence. For example, if new materials are developed by a materials science research group, they could increase the capabilities of a robotic gripper by allowing it to be equally strong but with less weight. Algorithms designed by computer scientists could greatly improve the predictive power used by autonomous vehicles.

Multidisciplinary centers that bring together contributing research labs can facilitate cooperation. Individual labs can be focused on specific areas and perform their research in the context of an overall robotics framework; the center itself can take the role of integrating technologies to see how they fit together.

Robotics research in academia is often coupled closely to industrial counterparts. The confluence is natural because research in robotics focuses so heavily on applications—and applications are the focus of industrial attention. In some subareas of robotics research, such as human-robot interactions, the sophistication of industrial activity outstrips that of academia. In some other areas, academic research pushes in directions that industry is not pursuing, such as theoretical aspects of robotics.

Multidisciplinary centers like the Stanford Robotics Center (SRC) are also useful for providing points of engagement with industry, pairing private companies with researchers on campus who have similar interests. In return for funding, corporate partners can more easily access new research and future employees.

RESEARCH FUNDING

Funding for robotics research comes from a variety of sources. The National Science Foundation supports foundational robotics research. Medical robots can be funded by the National Institutes of Health. Funding by the Defense Advanced Research Projects Agency (DARPA) is also common, with the Department of Defense sponsoring robot competitions to drive innovations they believe could be applicable to their use cases.

Another important source of research funding is the private sector. Tech companies like Google and Amazon want to know what is coming down the pipeline regarding robotics. They invest in academic labs and contribute to affiliate programs like those of Stanford's SystemX Alliance and the Stanford Robotics Center.

The amount of industrial funding devoted to robotics research and development far outstrips the amount available to academic researchers. Partly as a consequence, robotics faculty are being lured away from academia with compensation packages large enough to warrant concern at universities.[29]

NOTES

1. Ralph Lässig et al., "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. Statista, "Robotics—Worldwide," accessed August 30, 2023, https://www.statista.com/outlook/tmo/robotics/worldwide.

3. National Research Council, *Virtual Reality: Scientific and Technological Challenges* (Washington, DC: The National Academies Press, 1995), 313, https://doi.org/10.17226/4761.

4. Alan Winfield, *Robotics: A Very Short Introduction* (Oxford: Oxford University Press, 2012), 31, https://doi.org/10.1093/actrade/9780199695980.001.0001.

5. Winfield, *Robotics*, 20–22.

6. Winfield, *Robotics*, 34–35.

7. Mars Exploration Rovers, "Rover Update," National Aeronautics and Space Administration, accessed August 30, 2023, https://mars.nasa.gov/mer.

8. Bjorn Carey, "Maiden Voyage of Stanford's Humanoid Robotic Diver Recovers Treasures from King Louis XIV's Wrecked Flagship," Stanford News, April 27, 2016, https://news.stanford.edu/2016/04/27/robotic-diver-recovers-treasures.

9. "A Research Group in Carnegie Mellon University's Robotics Institute Is Creating the Next Generation of Explorers—Robots," *Science Daily*, July 19, 2023, https://www.sciencedaily.com/releases/2023/07/230719145936.htm.

10. National Science Foundation, "Award Abstract #1941933: CAREER: Fast, Furious and Fantastic Beasts: Integrative Principles, Biomechanics, and Physical Limits of Impulsive Motion in Ultrafast Organisms," accessed August 30, 2023, https://www.nsf.gov/awardsearch/showAward?AWD_ID=1941933.

11. Jacob Biba, "Microrobotics: Tiny Robots and Their Many Uses," Built In, September 6, 2022, https://builtin.com/robotics/microrobotics.

12. Stephen Gossett, "16 Agricultural Robots and Farm Robots You Should Know," Built In, March 7, 2023, https://builtin.com/robotics/farming-agricultural-robots.

13. "Robots in Disaster Response: Helping First Responders in Emergency," RoboticsBiz, June 16, 2023, https://roboticsbiz.com/robots-in-disaster-response-helping-first-responders-in-emergency.

14. Center for Sustainable Systems, "Autonomous Vehicles Factsheet," University of Michigan, September 2022, https://css.umich.edu/publications/facsheets/mobility/autonomous-vehicles-factsheet.

15. Guang-Zhong Yang et al., "The Grand Challenges of Science Robotics," *Science Robotics* 3, no. 14 (January 2018), https://doi.org/10.1126/scirobotics.aar7650.

16. "CHARM," Collaborative Haptics and Robotics in Medicine Lab, Stanford University, accessed August 30, 2023, https://charm.stanford.edu.

17. Alexander M. Kübler et al., "A Multi-Segment, Soft Growing Robot with Selective Steering," 2023 IEEE International Conference on Soft Robotics (RoboSoft), May 15, 2023, https://doi.org/10.1109/robosoft55895.2023.10122091.

18. Stanford Zhao Lab, "Soft Intelligence Materials Laboratory," Stanford University, accessed August 30, 2023, https://zhaolab.stanford.edu.

19. Stanford Autonomous Systems Laboratory, "Stanford ASL," Stanford University, accessed August 30, 2023, https://stanfordasl.github.io.

20. Stanford Biomechatronics Laboratory, "Designing Robots to Improve Mobility," Stanford University, accessed August 30, 2023, https://biomechatronics.stanford.edu.

21. Stanford Intelligent and Interactive Autonomous Systems Group, "Iliad," Stanford University, accessed August 30, 2023, https://iliad.stanford.edu.

22. Matt Travers and Howie Choset, "Bioinspired Robots: Examples and the State of the Art," American Association for the Advancement of Science, accessed August 30, 2023, https://www.science.org/content/page/bioinspired-robots-examples-and-state-art.

23. Wyss Institute for Biologically Inspired Engineering, "A Self-Organizing Thousand-Robot Swarm," Harvard University, August 14, 2014, https://wyss.harvard.edu/news/a-self-organizing-thousand-robot-swarm.

24. Christoph Bartneck et al., *Human-Robot Interaction: An Introduction* (Cambridge: Cambridge University Press, 2020), https://www.human-robot-interaction.org.

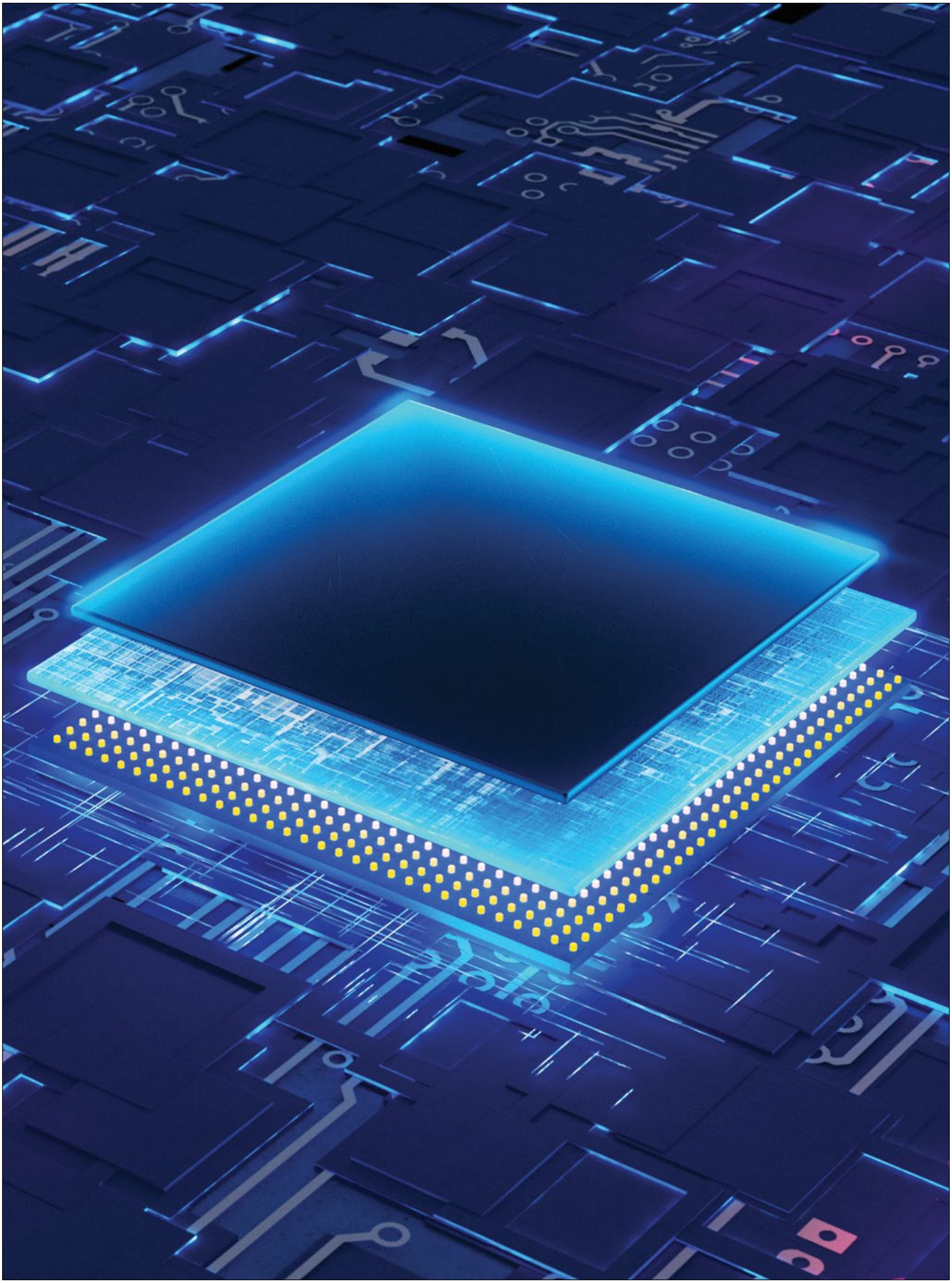
25. Artificial Intelligence Lab, "Stanford Robotics Lab," Stanford University, accessed August 30, 2023, https://cs.stanford.edu/group/manips/index.html; see also Oussama Khatib, "The Age of Human-Robot Collaboration," in *ROMANSY-22—Robot Design, Dynamics, and Control*, ed. Vigen Arakelian and Philippe Wenger (Rennes, France: Springer Cham, 2018).

26. Neil Savage, "Robots Rise to Meet the Challenge of Caring for Old People," *Nature*, January 19, 2022, https://www.nature.com/articles/d41586-022-00072-z.

27. James Wright, "Inside Japan's Long Experiment in Automatic Elder Care," *MIT Technology Review*, January 9, 2023, https://www.technologyreview.com/2023/01/09/1065135/japan-automating-eldercare-robots.

28. Gianmarco Veruggio and Fiorella Operto, "Roboethics: Social and Ethical Implications of Robotics," in *Springer Handbook of Robotics*, ed. Bruno Siciliano and Oussama Khatib (Berlin: Springer-Verlag, 2008), 1499–1524.

29. Cade Metz, "When the A.I. Professor Leaves, Students Suffer, Study Says," *New York Times*, September 6, 2019, https://www.nytimes.com/2019/09/06/technology/when-the-ai-professor-leaves-students-suffer-study-says.html.



08

SEMICONDUCTORS

KEY TAKEAWAYS

- Moore's law, which for fifty years has predicted rapid increases in semiconductor capabilities at decreasing costs, is now ending, raising profound implications for the future of hardware and software development.
- Recent research has identified methods that allow innovations in materials, devices, fabrication, and hardware to be added to existing processes or systems at low incremental cost. These methods need to be further developed since they will be essential to continue to improve the computing infrastructure we all depend on.
- Quantum computing may solve certain specialized problems, but experts debate whether it can ever achieve the rapid, consistent, predictable performance growth that semiconductors have enjoyed.

Overview

Semiconductors, or microchips, are crucial components used in everything from refrigerators and toys to smartphones, cars, computers, and fighter jets. These microchips 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 that penetrate or emanate from the human body. Precision robotic surgery would not be possible without digital-to-analog converter chips. Across these examples, innovation in chip design, new materials, and integration methods helps enable the performance, size, and efficiency of medical devices.[1]

A semiconductor is a material whose ability to conduct electricity lies between those of metals (with high electrical conductivity) and insulators (with low conductivity). Its conducting properties can be dramatically changed by adding impurities known as dopants into the material. The element silicon is the most commonly used semiconductor today, with the compound gallium nitride far behind in second place. This ability to change the state of the material allows one to build electrical devices such as transistors, which are the fundamental building block of all the information technology we use today.

Semiconductors are also used to create light-emitting diodes, or LEDs, for lighting and display applications, solar cells, electronics in power systems to control and manage the flow of electricity, and many other devices. This section focuses on semiconductors for computers and other information technology.

The first transistor was invented in 1949 and its size was measured in centimeters. Transistors were individual electronic devices that could be wired together along with other circuit components to do useful things, like function as a radio. About a decade later, the integrated circuit (IC), or chip, was invented; the IC integrates transistors and other circuit components into one physical package.

Different types of chips do different things in computers. Some chips are processors—they ingest digital data, perform computations on them, and output the results of those computations. Memory chips store information and are used with processors.

Many physical devices in daily life use chips, including computers, mobile phones and smartphones, cars, airplanes, washing machines, toasters, microwave ovens, televisions, refrigerators, lawn mowers, cameras, and so on. Chips control heating and cooling systems, elevators, fire suppression systems, and fire alarms in modern buildings. Traffic lights are controlled by chips. On farms, tractors and combines and irrigation systems are controlled by chips. Modern militaries could not function without chips in their weapons, navigation, and cockpit life support systems, or without computers to manage their complex logistics. The list goes on—in every aspect of modern life, chips are essential.

Two aspects of chips are important for our purposes here. Chips must be designed and then manufactured, both of which call 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 has a physical component that calls for large factories, or “fab facilities,” that can produce chips by the millions and billions. In 2022, a record 1.15 trillion chips were shipped worldwide.[2] But fabrication also entails a substantial degree of process engineering to continue to improve process technology and to achieve the stringent manufacturing standards. For example, the “clean rooms” in which chips are made require air that is one thousand times more particle-free than what is in a hospital operating room.[3]

Because chip design and chip fabrication are so different in character, only a very few companies do both. Intel is one such company. Many companies specialize in design—Qualcomm, Broadcom, Apple, and Nvidia. Such companies are also called “fabless” in recognition of the fact that they do the design work and outsource fabrication to others on the theory that the former has relatively high market value and the latter has relatively low market value.

Today, “others” usually means one company: Taiwan Semiconductor Manufacturing Company (TSMC), which is by far the world’s largest contract chip manufacturing company. Samsung in South Korea is a distant second, and the United Microelectronics Corporation (UMC), also in Taiwan, ranks third. With a large fraction of the world’s chip-manufacturing capacity in Taiwan, the global supply chain for chips is clearly fragile. US manufacturing capacity is losing ground. Fabrication plants in the United States accounted for 37 percent of global production in 1990 but dropped to just 12 percent by 2021.[4]

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, 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. As such, it depends not only on the sophistication and capabilities of the equipment and facilities used to fabricate chips but also on a set of economic conditions and decisions that make it financially sensible to invest in the expensive construction of new state-of-the-art fabrication facilities. Including the research and development needed to develop appropriate manufacturing processes, the cost of such a facility may be as much as $20 billion.

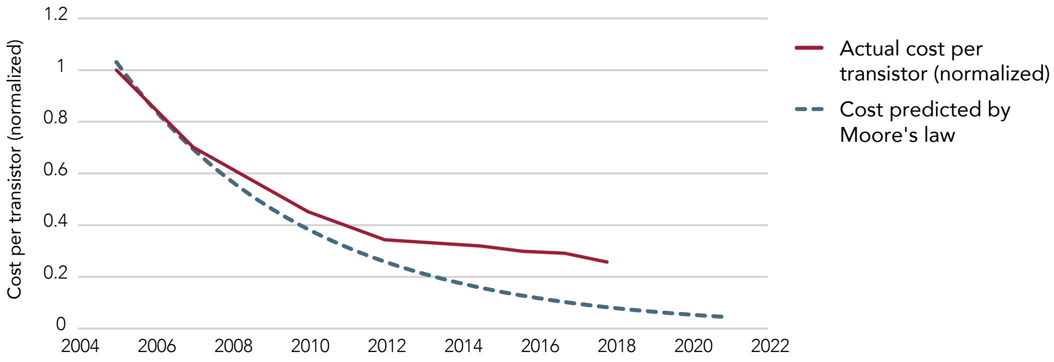
Although Moore's law is often stated as the number of transistors doubling every few years, historically the cost of making a chip was mostly independent of what it holds. This means that every few years, the same-size and approximately same-cost chip will have twice the number of transistors on it.

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 a more powerful system for the same cost. This scaling has been so consistent that everyone expects the cost of computing to decrease with time. The expectation is so pervasive that in almost all fields of work, people are working to develop more complex algorithms to achieve better results while relying on Moore's law to rescue them from the consequences of that additional complexity.

Unfortunately, the end of Moore's law cost scaling appears to be in sight, and the end of that technology trend in chip fabrication has profound implications for future systems and design.

As figure 8.1 shows, the actual cost of a chip per transistor (as shown by the solid red line) was tracking the cost predicted by Moore's law (dashed blue line) relatively well from 2004 to 2012.[5] However, the actual

FIGURE 8.1 Cost per transistor over time



Source: Data from Qualcomm Technologies Inc., https://eri-summit.darpa.mil/docs/Mollenkopf_Steve_Qualcomm_Final.pdf cost per transistor started to level off around 2012, and it has not kept up with Moore's law predictions since then.[6] Similar slowdowns in scaling the cost per bit of dynamic random-access memory (DRAM) and the cost per bit of mass disk memory have occurred as well. The cost per bit of flash memory has continued to drop, though not as rapidly as before, and it too may continue to slow.

Quantum Computing

While tremendous effort has been and is being invested to continue scaling semiconductor technology, the benefits gained from this scaling have dramatically decreased. This slowing has increased interest in alternative technologies that might have potential advantages over today's semiconductors. One such technology is quantum computing.

Quantum computing (QC) is a different way of performing computation—this is both its allure and its difficulty. Its different framing allows some computational tasks to take fewer steps than on today's computers, potentially making it much faster. Unfortunately, the different framing also means that the basic operations are different and more noise sensitive, which requires completely different hardware, and solving problems requires a different approach to algorithm design. Therefore, creating a useful quantum computer requires innovation across the entire quantum software/hardware stack: algorithms, compilers, control electronics, error correction, and quantum hardware.

A classical computer uses individual bits as the smallest unit of data, each being 0 or 1, and represents numbers, words, and even pictures as larger collections of bits. Analogously, QC uses quantum bits, or "qubits." Unlike normal bits, quantum mechanics allows a qubit to be in many different states beyond just 0 or 1. It can be both a 0 and a 1 at the same time, for example, allowing a quantum computer to process a large number of possibilities and problems all at once. Much has been written about this "quantum parallelism," but building algorithms that make use of it is difficult.

Qubits need to be isolated quantum systems, and many technologies are under consideration for the physical construction of qubits; these include trapped ion, superconducting, cold atom, photonic, crystal defect, quantum dot, and topological technologies. The most advanced quantum computing machines use either trapped ion or superconducting qubits, but neither technology has a clean scaling path to larger machines—so work continues on other possible technologies.

Current hardware is focused on increasing the number of qubits while decreasing the number of errors. The best quantum machines today have around fifty qubits and can do about three hundred 2-qubit operations between errors. Historically, increasing qubits increases error rates: machines with four hundred qubits exist but can do fewer than fifty operations. For comparison, conventional computers have billions of bits and can do more than a million billion (10^{15}) operations before errors occur.

Given the higher error rate, quantum error correction algorithms have been developed to work using noisy physical qubits. The essential idea is to encode a logical qubit among multiple physical qubits—maybe as few as ten, maybe as many as one thousand—and take advantage of the ensuing redundancy to detect and correct errors in the physical qubits so the logical qubits remain correct. Algorithms with smaller overheads require lower base error rates to operate correctly, and today's machines are approaching the point where error correction can be demonstrated.

For QC to be successful, it will be necessary to scale the number of qubits that can be used in a machine and decrease the error rates they experience. Like all technological development, this scaling will take greater financial investment. Moore's law was the result of a virtuous cycle where better technology increased the semiconductor market revenue, which in turn increased funding available to develop the next technology generation. For QC to flourish, it too will need a virtuous cycle created by a growing market that funds increasingly difficult technology development.

How to create a growing market for quantum computers is one of the biggest challenges in the field. While in theory a quantum computer can solve some problems—such as finding the prime factors of a number much faster than a conventional computer—solving this problem would require a computer that is much more advanced than what we can build today. To be able to refine the technology to that point, we first need to produce smaller useful machines that can generate the funding needed for continued technology development.

As of today, no one has found a commercial problem that a near-term quantum computer can solve that can't also be solved as effectively on a conventional computer. Given the large initial cost of a quantum computer relative to conventional computing, this situation means that we still don't have a commercial market for quantum computing. Factoring is still the holy grail of quantum computing, but the rough consensus in the field is that it will be at least a decade before it is possible to use quantum computing to factor large numbers—a problem that underlies the security offered by much of today's public-key cryptography.

Over the Horizon

Impact of New Technologies

The timeline for quantum computing is uncertain. But even if and when it does arrive and quantum computing is fully successful, QC machines will be useful for only a limited class of applications; they won't replace today's semiconductor technology.

If it is true that Moore's law is at an end, improvements in end-user applications will come primarily from better optimization of algorithm/hardware/technology to the application, rather than technology scaling. Such optimization will require innovation across the entire technology stack, in new materials, new technologies and devices, and design methods. Yet finding these new methods of building systems poses a dilemma for the industry that in some ways is burdened by its own success.

Moore's law has enabled us to produce computing systems of amazing power and complexity, but they also require huge and expensive design teams and must be manufactured in fabs costing billions of dollars to construct. As a result, the industry has consolidated, and the number of companies working in the hardware space has diminished, as has student interest.

The result is a paradox. Performing the necessary optimizations requires innovative researchers willing to try radical ideas that in the end might not succeed. But how can we find researchers and companies willing to take on these risks if every attempt costs $100 million or takes two years?

We must make it easier and cheaper and faster for people to explore those innovations. Many research teams are working on this problem, and it is an active research topic at Stanford. It calls for making the complexity of the design tasks proportional to the change being made, rather than to the complexity of the resulting system. Imagine how little home remodeling would be done, for example, if every idea for remodeling entailed revision of all the blueprints for the entire house, as though everything had to be redone from scratch. The latter more closely corresponds to the process of state-of-the-art chip design today.

The goal is to allow the prototyping of solutions at low cost. The specific tools and approach to accomplish this goal depend on which level in the semiconductor design stack is being addressed. For example, to foster hardware optimization to help with specific software applications, the designers of those software applications need to be able to explore different ways in which hardware improvements would help them to optimize their applications.

Software application developers should be able to test various custom accelerators without requiring a deep knowledge of hardware design. Because their primary concern is the potential benefits a new hardware accelerator could offer their applications, they should be able to direct the hardware design by exploring different structures and configurations in their code. This whole approach relies on the existence of a base hardware platform on which the application currently runs. Stanford hardware researchers interested in this approach are exploring tools and new software and hardware interfaces to the base platform that would allow an application developer to make a small addition to the platform to improve the application's performance. The goal is to build an infrastructure that allows application designers to add small hardware extensions to the base platform. This hardware extension is used only for the specific application of the software developer. In essence, just as a housing developer can build standard houses but also allow buyers to customize a room, this approach would enable software application developers to customize hardware for their applications by making a small addition to a standard base platform rather than starting from scratch.

Of course, this model works only if there are available platforms to use. Since building a platform is very expensive, for this idea to be a success it is critical to convince some firms that have complete working systems—for example, Nvidia, Apple, Intel, AMD, and Qualcomm—to participate in the effort. This approach bears substantial similarity to the model of the app store, which provides an open interface while keeping the base system proprietary. The app store model balances open innovation with the profit motives of companies. Advanced packaging technology makes this approach possible, yet many economic issues still need to be resolved.

Challenges of Innovation and Implementation

Historically, the US government has tended to refrain from funding R&D efforts that it believes are more properly supported by the private sector. These efforts are characterized by having relatively short time scales to benefit the specific competitive strategies that a private company might have. By contrast, it has been more generous in funding academic research with long time horizons on the grounds that such research is precompetitive in nature and enables "all boats to rise with the tide."

There are good theoretical, science-based arguments for why quantum computing is feasible in principle. It is further known that some types of useful problems are indeed solvable by quantum computing algorithms.

But there are two potential flies in the ointment. First, though QC may be feasible in principle, a large amount of engineering stands in the way of making QC feasible in practice. Second, the problem of QC's utility is not a matter of whether it can solve certain useful problems, but whether it can solve them more rapidly than conventional computers. This latter point means that QC's value depends not just on the status of the field of QC but also on the status of its competitors.

In any event, the private sector is unlikely to fund QC research at the level necessary to determine its actual utility because of the long time horizons to any possible payoff. Thus, if QC research is to flourish, sustained US government support will be necessary.

Policy, Legal, and Regulatory Issues

GEOPOLITICS

Taiwan is where over 60 percent of the world's semiconductor chips and over 90 percent of the most advanced chips are fabricated, most by Taiwan Semiconductor Manufacturing Corporation. Both China and the United States, as well as the rest of the world, depend on TSMC for the chips that power their advanced technologies.[7]

The fragility of the supply chain for advanced chips was demonstrated with the outbreak of the COVID-19 pandemic and its associated lockdowns. As of April 2021, more than 169 different industries were impacted by the lack of supply of semiconductors, according to an analysis by Goldman Sachs.[8] Given the stated desire of China to reintegrate Taiwan into the People's Republic of China—and its refusal to disavow the use of force to do so—the stakes for the economy and national security of the United States could not be higher. Few scenarios are more dire than physical PRC control over Taiwan and PRC global dominance in the manufacture of semiconductor chips.

In response to these concerns, the United States passed the CHIPS Act in 2022, which aims to increase domestic semiconductor production and research.[9] The act appropriated $52.7 billion for semiconductor manufacturing, research, and workforce development, along with a 25 percent tax credit for private investments made for capital expenses related to the manufacture of semiconductors—with an estimated value of $24.5 billion.[10]

The act was meant to provide incentives for companies to invest in American fabs. Companies have indeed announced $166 billion in investments in the one year since the law was enacted, though many of those projects are contingent on approval of federal aid.[11]

In addition, the US government in October 2022 imposed tighter export control restrictions on various semiconductor-related products that might otherwise be headed for China.[12] These products include advanced computing chips—such as the Nvidia chips used for AI computing—and equipment that could help to manufacture advanced semiconductor chips. It also reportedly persuaded nations such as Japan and the Netherlands to adopt similar export control measures, thus imposing a near-total blockage of China's ability to buy the equipment necessary to make leading-edge chips.[13] National Security Advisor Jake Sullivan noted that the rule shifted US export policy: "We previously maintained a 'sliding scale' approach that said we need to stay only a couple of generations ahead. That is not the strategic environment we are in today. Given the foundational nature of certain technologies, such as advanced logic and memory chips, we must maintain as large of a lead as possible."[14]

It is worth noting that the ending of Moore's law complicates this logic. Moore's law (i.e., the continued decrease in computational costs) is nearing its end because of limitations imposed by the laws of physics, which apply to all nations. This has several implications. First, pouring more money into chip fabrication won't increase transistor density forever, and Western nations will hit these limits sooner than nations like China because the West is currently ahead. Second, while the West may have previously benefited from restricting China's access to advanced fabrication technology, these controls are less impactful now. If newer technology offers diminishing returns, then older technology will not lag that far behind. Consequently, the Western lead will diminish over time. Export controls that limit China's access to Western technology also incentivize China to develop indigenous expertise and boost its domestic industry by providing a larger market for Chinese firms. Additionally, export controls can hamper US academic semiconductor research, potentially stalling innovation by limiting access for international talents.

Pouring more money into chip fabrication won't increase transistor density forever.

TALENT

As noted above, regular hardware-based improvement for complex applications is unlikely to be sustained by Moore's law for much longer, if at all. This reality puts a premium on generating new ideas for hardware improvements that are application specific. One new approach to integrate hardware improvements into software applications was described above, but equally important is the availability of human talent.

However, student interest in hardware design has dropped precipitously, as technically inclined students tend by overwhelming margins to favor software-oriented jobs. Some estimates suggest that by 2030, the semiconductor manufacturing employment sector will be able to fill only 30 percent of its needs.[15] Since appropriately trained people are the only real source of new ideas, these trends do not bode well for the industry.

As one possible data point, it is noteworthy that the CEO of TSMC pointed to "an insufficient amount of skilled workers with the specialized expertise required for equipment installation in a semiconductor grade facility" as an important reason that the construction of a planned fabrication facility in Arizona was significantly behind schedule.[16] To be fair, it is not clear whether he was referring to the high-end chip designers; others have suggested that TSMC's difficulties stem from clashes of work culture between Taiwan and US unions that oppose bringing in nonunion workers from Taiwan to help build the plant.[17]

CLASSIFICATION

In 2014, Edward Snowden leaked a US classified document pointing to the existence of an $80 million classified program in quantum computing, apparently for the purpose of codebreaking.[18] The scope and nature of classified US government programs currently pursuing quantum computing are unknown. Regardless, a robust unclassified program to support QC research will help to ensure that quantum scientists in the United States are able to draw on ideas from around the world—and maintain our leadership in the field.

NOTES

1. Semiconductor Industry Association, *From Microchips to Medical Devices: Semiconductors as an Essential Industry during the COVID-19 Pandemic*, 2020, https://www.semiconductors.org/wp-content/uploads/2020/10/From-Microchips-to-Medical-Devices-SIA-White-Paper.pdf.

2. Arjun Kharpal, "Global Semiconductor Sales Top Half a Trillion Dollars for First Time as Chip Production Gets Boost," CNBC, February 15, 2022, https://www.cnbc.com/2022/02/15/global-chip-sales-in-2021-top-half-a-trillion-dollars-for-first-time.html.

3. HPCwire, "Inside an Intel Chip Fab: One of the Cleanest Conference Rooms on Earth," March 29, 2018, https://www.hpcwire.com/off-the-wire/inside-an-intel-chip-fab-one-of-the-cleanest-conference-rooms-on-earth.

4. Semiconductor Industry Association, *2021 State of the U.S. Semiconductor Industry*, 2021, https://www.semiconductors.org/wp-content/uploads/2021/09/2021-SIA-State-of-the-Industry-Report.pdf.

5. Electronics Resurgence Initiative 2.0, "2019 Summit Agenda, Videos, and Slides," Defense Advanced Research Projects Agency, accessed August 30, 2023, https://eri-summit.darpa.mil/2019-archive-keynote-slides.

6. In fact, transistor costs are usually plotted in a semilog 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.

7. *The Economist*, "Taiwan's Dominance of the Chip Industry Makes It More Important," March 6, 2023, https://www-economist-com/special-report/2023/03/06taiwans-dominance-of-the-chip-industry-makes-it-more-important.

8. Allison Nathan, "The Global Chip Shortage: Impact, Outlook and Recovery," June 29, 2021, in *Exchanges at Goldman Sachs*, podcast produced by Goldman Sachs, MP3 Audio, 19:40, https://www.goldmansachs.com/intelligence/podcasts/episodes/06-29-2021-hari-yuzawa-hall.html.

9. Semiconductor Industry Association, "CHIPS for America Act and FABS Act," accessed August 30, 2023, https://www.semiconductors.org/chips.

10. Congressional Research Service, *Frequently Asked Questions: CHIPS Act of 2022 Provisions and Implementation*, April 25, 2023, https://crsreports.congress.gov/product/pdf/R/R47523.

11. White House, "Fact Sheet: One Year after the CHIPS and Science Act, Biden-Harris Administration Marks Historic Progress in Bringing Semiconductor Supply Chains Home, Supporting Innovation, and Protecting National Security," Statements and Releases, August 9, 2023, https://www.whitehouse.gov/briefing-room/statements-releases/2023/08/09/fact-sheet-one-year-after-the-chips-and-science-act-biden-harris-administration-marks-historic-progress-in-bringing-semiconductor-supply-chains-home-supporting-innovation-and-protecting-national-s.

12. Bureau of Industry and Security, "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)," US Department of Commerce, October 7, 2022, https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file.

13. "The United States Announces Export Controls to Restrict China's Ability to Purchase and Manufacture High-End Chips," *American Journal of International Law* 117, no. 1 (January 2023): 144–50, https://doi.org/10.1017/ajil.2022.89; Yuka Hayashi and Vivian Salama, "Japan, Netherlands Agree to Limit Exports of Chip-Making Equipment to China," *Wall Street Journal*, January 28, 2023, https://www.wsj.com/articles/japan-netherlands-agree-to-limit-exports-of-chip-making-equipment-to-china-11674952328.

14. White House, "Remarks by National Security Advisor Jake Sullivan at the Special Competitive Studies Project Global Emerging Technologies Summit," Speeches and Remarks, September 16, 2022, https://www.whitehouse.gov/briefing-room/speeches-remarks/2022/09/16/remarks-by-national-security-advisor-jake-sullivan-at-the-special-competitive-studies-project-global-emerging-technologies-summit.

15. Tom Dillinger, "A Crisis in Engineering Education—Where Are the Microelectronic Engineers?," SemiWiki, July 3, 2022, https://semiwiki.com/events/314964-a-crisis-in-engineering-education-where-are-the-microelectronics-engineers; Dylan Martin, "US Chip Industry Has Another Potential Shortage: Electronics Engineers," *The Register*, July 8, 2022, https://www.theregister.com/2022/07/08/semiconductor_engineer_shortage.

16. Tobias Mann, "TSMC Says Arizona Fab behind Schedule, Blames Chip Geek Shortage," *The Register*, July 20, 2023, https://www.theregister.com/2023/07/20/tsmc_arizona_fab.

17. Jacob Zinkula, "TSMC's Arizona Plant Is Delayed Over Poor Management, Not a Shortage of US Skilled Labor, the Workers Building It Say," Business Insider, August 28, 2023, https://www.businessinsider.com/tsmc-phoenix-arizona-chip-factory-taiwan-semiconductor-management-safety-workers-2023-8; Kevin O'Marah, "TSMC vs. Arizona Pipe Trades 469 Union: Everyone Loses," *Forbes*, August 17, 2023, https://www.forbes.com/sites/kevinomarah/2023/08/17/tsmc-vs-arizona-pipe-trades-469-union-everyone-loses.

18. Steven Rich and Barton Gellman, "NSA Seeks to Build Quantum Computer That Could Crack Most Types of Encryption," *Washington Post*, January 2, 2014, https://www.washingtonpost.com/world/national-security/nsa-seeks-to-build-quantum-computer-that-could-crack-most-types-of-encryption/2014/01/02/8fff297e-7195-11e3-8def-a33011492df2_story.html.

09

SPACE

KEY TAKEAWAYS

- Space technologies are increasingly critical to everyday life (e.g., GPS navigation, banking, missile defense, internet access, and remote sensing).
- Space is a finite planetary resource. Dramatic increases in satellites, debris, and competition are threatening access to this global commons.
- Private-sector actors play a critical and growing role in many aspects of space-based activities (e.g., launch, vehicles, and communications), because they offer better, cheaper, and rapidly deployable capabilities.

Overview

Sputnik 1 was the world's first artificial satellite, placed into orbit by the Soviet Union in 1957. A technology demonstration, Sputnik broadcast an easily monitored radio signal from space for a few weeks. This little 184-pound, 2-foot-diameter capsule launched the space age—and today many thousands of satellites provide Earth-bound nations and their citizens with communications, navigation, multispectral observation, and imagery of terrestrial phenomena that are useful in many walks of life. A substantial amount of scientific discovery is also made possible with space-borne instrumentation. Finally, space operations support military forces on Earth and thus space itself is a domain in which international conflict and competition play out.

The global space sector experienced growth in 2022, driven primarily by commercial and private activities.[1]

It was valued at $424 billion in 2022, showing an 8 percent growth from the previous year, with projections suggesting it might reach $737 billion within the next decade.[2] This growth coincided with a shift from government-operated launches to private providers. There were 180 global rocket launches in 2022, an increase from the previous year.[3] Private space investments in 2022 saw a 25 percent reduction, attributed to an economic downturn affecting start-ups, though the year maintained high investment figures.

By definition, space technology is any technology developed for the purpose of conducting or supporting activities beyond the Kármán line (i.e., 100 kilometers or 62 miles above Earth's surface). A space mission is a system of systems that is designed to optimally accomplish objectives, generally through the art and science of space systems engineering. A space mission includes several components:

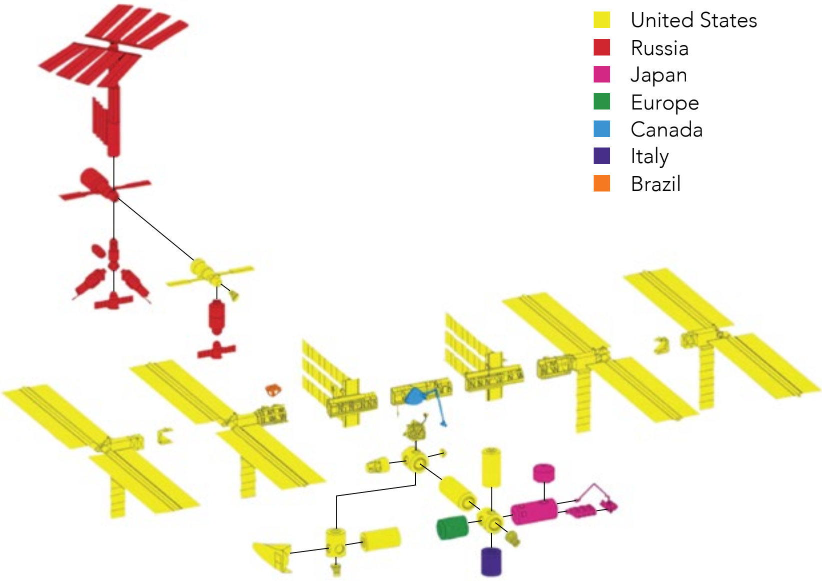
- The mission objectives, which can be scientific, commercial, or military

- A space segment, which includes the spacecraft and the orbits that have been selected and designed to accomplish the objective of the space mission

- A ground segment, which includes the rocket launcher, ground stations, and mission control centers

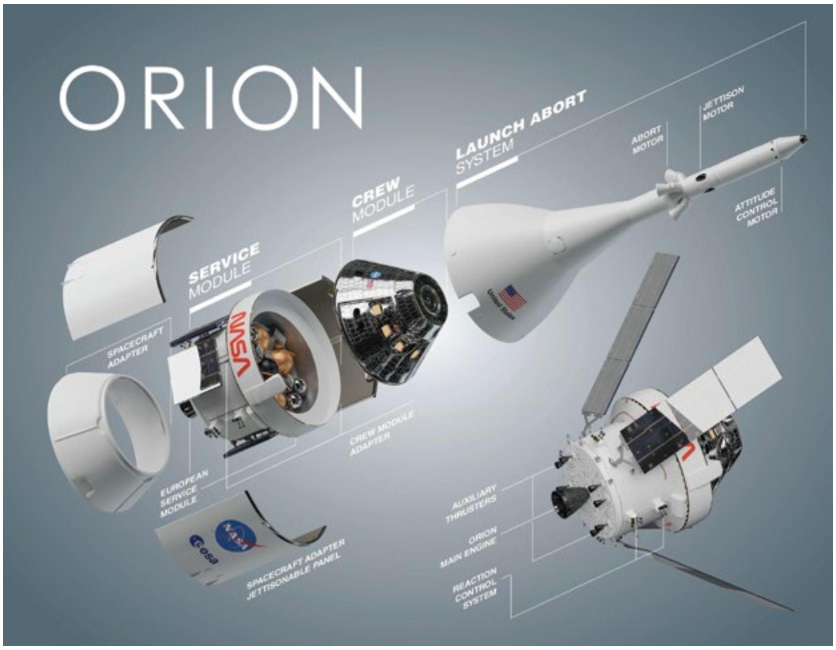
An example of an extremely complex space mission is the International Space Station (ISS) (see figure 9.1). The space segment includes the ISS itself, the structure that is the space station, several cargo and crew vehicles (e.g., SpaceX Dragon, Soyuz, Automated Transfer Vehicle), and data relay satellites (e.g., the Tracking and Data Relay Satellite system). The ground segment includes the rockets

FIGURE 9.1 The International Space Station



Source: National Air and Space Museum Archives via Smithsonian Institution

FIGURE 9.2 NASA's Orion spacecraft



Source: NASA

that are used to deploy those elements into space as well as a worldwide network of ground-based control centers.

A key part of a space mission is a spacecraft consisting of a payload and a bus (see figure 9.2). The payload is a collection of instruments that are used to achieve the mission objectives, while the bus is anything else that supports the payload in achieving those objectives. The bus consists of a variety of subsystems that are analogous to individual organs within a living being. These subsystems manage (1) translational (orbit) and rotational (attitude) motion of a spacecraft; (2) communications satellite-to-ground and satellite-to-satellite when available; (3) data storage and processing; (4) generation, distribution, and dissipation of electrical power to all other subsystems and to the payload as well; (5) thermal control to ensure that the components of the spacecraft are within their maximum and minimum temperature operational and survival limits; and (6) the structure to hold the various components and protect them from the physical stresses encountered during the mission lifetime.

One way to classify space systems is by whether they are crewed or uncrewed. The former includes systems used for crew transportation to and from space (Orion, Soyuz, Dragon), space stations (ISS, Lunar Gateway, Tiangong, Orbital Reef), and land-based surface systems to provide human habitat, in situ resource utilization, and scientific experimentation. The first crewed space flight from US soil since 2011 was a suborbital test flight operated by Virgin Galactic, a private company. Crewed US access to the International Space Station since 2011 has been aboard rockets operated by Russia and more recently by the private company SpaceX. The NASA-operated Artemis program plans to launch its first crewed mission, a moon flyby, in late 2024.

Uncrewed systems include systems for Earth and planetary remote sensing (the Gravity Recovery and Climate Experiment; Doves, SkySats, and RapidEye of the private company Planet; the Mars Reconnaissance Orbiter); communication and navigation (the Tracking and Data Relay Satellite system, the European Galileo, Starlink); astronomy and astrophysics (the Hubble, James Webb, and Nancy Grace Roman telescopes); space logistics/in-space assembly and manufacturing (Restore-L, the Mission Extension Vehicle); and planetary exploration (Perseverance, Ingenuity, Zhurong).

Alternatively, space systems can be characterized by size. At one end of the distribution are large structures such as the International Space Station with a mass of about 420 tons and a truss length of 94 meters. New commercial space stations include Orbital Reef, being built by Jeff Bezos's Blue Origin; Starlab by Airbus; and an as-yet unnamed space station by Northrop Grumman. These commercial space stations are set to eventually replace the outdated ISS, which is expected to be retired in 2030, and they will be of comparable size.

At the other end of the distribution are much smaller satellites, often called smallsats. A NASA Ames Research Center report classifies anything under 500 kilograms as a small spacecraft.[4] This includes Sprite chipsats, which weigh less than a gram and were developed by a former faculty member of the Stanford Aeronautics and Astronautics department. CubeSats are the most popular small satellites today. Introduced in 1999 by Bob Twiggs of the Stanford Aeronautics and Astronautics department, each CubeSat unit measures 10 x 10 x 10 cm, weighs a kilogram, and can be combined to build larger satellites. Originally intended for educational use, CubeSats now support a growing commercial market. Today, a large majority of functional satellites in space weigh between 100 and 1,000 kilograms.

Space systems can be characterized by the orbits where they move. One category refers to objects in orbit around Earth: such space systems can be classified as being in low Earth orbit (LEO, less than 1,000 km in altitude), medium Earth orbit (MEO, between 2,000 and 35,000 km in altitude), high elliptical orbit (HEO), or geosynchronous orbit (GEO). The image in figure 9.3 was obtained by a commercial satellite in LEO, operated by a space start-up founded by Stanford students.

FIGURE 9.3 Perth, Australia, as seen from SkySat-1

Source: Planet Labs PBC

Another category where space systems move is defined by the Lagrange points in space, the most significant of which are hundreds of thousands or a million kilometers away. Lagrange points are defined with respect to two bodies, such as the Sun and Earth or the Earth and Moon, and denote points around which an orbiting spacecraft can remain in a fixed spatial relationship to the two considered bodies.

The most well-known Lagrange point orbits are referred to as halo orbits and were discovered at Stanford in 1966 by Professor John V. Breakwell and his PhD student Robert Farquhar. Lagrange point orbits provide many benefits; for example, in a Sun-Earth halo orbit, the fixed spatial relationship of a spacecraft relative to the Earth and Sun means that it is possible to view Earth with the same illumination conditions—that is, with sunlight shining on terrestrial objects with the same intensity, from the same angles, and casting the same shadows. Furthermore, Earth-Moon Lagrange points are particularly good places in space at which to stage missions entailing travel between the Earth and Moon.

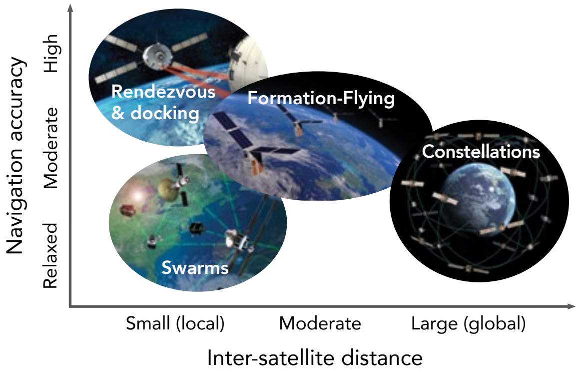
Finally, a number of interplanetary probes have been launched to every planet—and some asteroids—in our solar system. Some have gone beyond the solar system, including Voyager 1 and Voyager 2, which were both launched in 1977 and are now the human-made objects most distant from Earth today.

The space systems described above generally consist of a single spacecraft, which has limited capability because of its constraints in carrying capacity (size and volume) and maneuverability (propellant). Distributed space systems—made of two or more spacecraft that interact and sometimes work together—can accomplish objectives that would otherwise be very difficult or impossible with a single spacecraft.

A distributed system can be characterized along two axes—the distance between space segment components, which could be virtually nothing to tens of thousands of kilometers, and the positional accuracy needed of each component relative to other components (see figure 9.4). In this classification, several architectures emerge:

- **Rendezvous and docking** are characterized by small separations and high positional accuracy. This architecture was necessary for the United States to land humans on the moon and today is a key technology needed for removal of space debris from orbit, for in-orbit servicing of satellites, and to assemble and manufacture larger structures in space.
- **Formation-flying** architectures are needed for observational missions that call for large effective apertures, such as space-based telescopes whose optical components are controlled very precisely with respect to one another at separations of tens to hundreds of meters. Gravimetric and interferometric missions require the same architecture.
- **Swarms** that sense the environment or share resources such as power or computation remotely also need to be kept in a relatively tight formation (i.e., close to each other), but the components do not necessarily need to be at fixed distances from one another.
- **Constellations** have components that are separated by tens of thousands of kilometers for global ground coverage, but their relative positioning need not be particularly precise. Examples of constellations include satellites of the Global Navigation Satellite System (GNSS) such as the US GPS, the Chinese BeiDou, the Russian Globalnaya Navigatsionnaya Sputnik Sistema (GLONASS), and the European Galileo systems; communications satellites providing worldwide coverage such as Starlink; and remote sensing and imaging satellites.

FIGURE 9.4 Characterizing a distributed space system



Source: Diagram by Simone D'Amico from NASA images

Key Developments

Space technology has proved its value to the national interest. Some of the most important applications today include:

Navigation This includes, more generally, position, navigation, and timing services around the world and in space. GPS satellites (and those of other nations as well) help people know where they are and how fast and in which direction they are going, whether they are 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 is available anywhere in the world. This is a key tool for the financial sector, electric power grid, and transportation.

Communications Although the vast bulk of international and long-haul communications traffic is still routed through landlines (mostly fiber optic), satellites provide voice and data communications as well as internet access in otherwise inaccessible places around the world and, of course, for mobile phone users in cars and planes and on ships. Recent innovations in space-based communications technology include the development of optical communication systems—which use light to carry data and offer higher bandwidth and security. These include laser communications both space-to-ground and space-to-space, which hold a particularly high value for government and military.

Remote sensing Satellites gather information about a geographical area, the environment, or an object by detecting and measuring energy that may be reflected or emitted by the entity being sensed. These satellite systems generate data to create a "digital twin of the Earth" for disaster prediction, prevention, monitoring, mitigation, and recovery, and will play a huge role in the future by enabling simulation and prediction of terrestrial phenomena and especially disasters. Space-based remote sensing is used to observe and surveil large forest fires, weather formation, the evolution of cloud cover, erupting volcanoes, dust storms, changes in the geography of a city or in farmland or forests (e.g., as the result of fires, earthquakes, or flooding), changes in terrain such as glacier movement or landslides, and surface topography. Space-based remote sensing can scan large areas of Earth rapidly, though at some cost in resolution. Remote sensing also varies by revisit rate—revisiting on the order of hours or days is needed for rapidly unfolding phenomena, such as the progression of a hurricane, while applications such as glacier monitoring require much less frequent measurements.

Scientific research Space-based telescopes, such as the James Webb space telescope, play an important role in various areas of astronomy and cosmology. They help in studying the earliest stars and understanding the creation of the first galaxies and offer in-depth insights into the atmospheres of planets that might support life.

Space transportation The space transportation industry is becoming increasingly privatized and provides launch services for parties wanting to orbit satellites and transport services to in-orbit space stations. The costs of placing payloads into LEO have fallen from a high of $65,000 per kilogram to $1,500 per kilogram in 2021,[5] largely driven by the advent of multiple launch capability of a single rocket—as many as 100 to 150 at a time—coupled with reusable rocket launch vehicles.

National security Space-based satellites scan Earth looking for launches of ballistic missiles that may be aimed at the United States or its allies, for nuclear weapons explosions on the surface anywhere in the world, and for radio traffic and radar signals from other countries. Of course, all these applications—navigation, communications, and remote sensing—are valuable in a military context.

There is an increasing trend toward privatization across most space technologies as the space sector moves away from legacy space technologies owned by the government or large contractors. These legacy systems are characterized by large, expensive spacecraft with long development timelines. Today, a "NewSpace" economy is turning to private companies, creating a global space environment in which systems and services are more accessible and

> Satellite systems generate data to create a "digital twin of the Earth" . . . and will play a huge role in the future by enabling simulation and prediction of terrestrial phenomena and especially disasters.

less expensive—and available to all. Governments are also looking to commercial space for new capabilities, like awarding contracts to private companies to develop smallsat constellations or for in-space servicing, assembly, and manufacturing (ISAM).

Over the Horizon

Impact of Space Technologies

Future applications of space technology are likely to include:

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, the microgravity environment of space means that the effects of gravity on fabrication can be minimized, enabling more perfect crystals and more perfect shapes to be fabricated, to give examples.

Mining The moon and asteroids may well have vast storehouses of useful minerals that are hard to find or extract on Earth (e.g., rare-earth elements). Future space mining operations may bring some of these to Earth or utilize them for further human expansion in the solar system.

Power generation It is well known that the sun's radiation on Earth can generate electricity through solar cells. But above Earth's atmosphere in certain orbits, the sun never stops shining; indeed, it shines more brightly because it is not attenuated by Earth's atmosphere or by weather. It may be economically feasible in the future to capture such energy and beam it to Earth for sustainable electrical generation.

National security Although the Outer Space Treaty prohibits the placement of nuclear weapons or other weapons of mass destruction in space, there are no restrictions on other military uses of space, including the placement of conventional weapons in space. Furthermore, space-based capabilities are integral to supporting modern warfighters; accordingly, they will be the targets of foreign counterspace threats. Rapid-launch capabilities to facilitate fast replacement of satellites rendered inoperative during times of war or conflict will increase the resiliency of critical national space assets.

In-space logistics, servicing assembly, and manufacturing (ISAM) Dominance, security, and sustainability in space require infrastructure that supports cheap, quick, reliable access and the ISAM capabilities to approach, inspect, assess damage to, repair, prolong the lifetime of, retire, or remove space assets without jeopardizing the space environment.[6] Spacecraft autonomy, in combination with Rendezvous, Proximity Operations, and Docking (RPOD), is a critical technology for ISAM. For example, orbital tugs, in-orbit fuel depots, or orbital transfer vehicles (OTVs) are needed for space logistics and to enable a circular space economy.

Challenges of Innovation and Implementation

Public entities, driven by the need for public accountability, have become more risk averse, often showing reluctance to embrace innovation unless traditional methods are unviable. In contrast, private companies pursue innovation when it provides economic viability and a competitive edge via intellectual property. Collaborative efforts between academia and industry are pivotal for technology commercialization and real-world demonstrations of advancements codeveloped by industry and academic partners.

The emergence of low-cost, high-quality information from space-based assets—increasingly launched and operated by private companies—is also an important driver of open-source intelligence (OSINT) that data analysts can buy on the open market. OSINT threatens to upend traditional intelligence gathering as closely held information and analysis becomes more readily available.

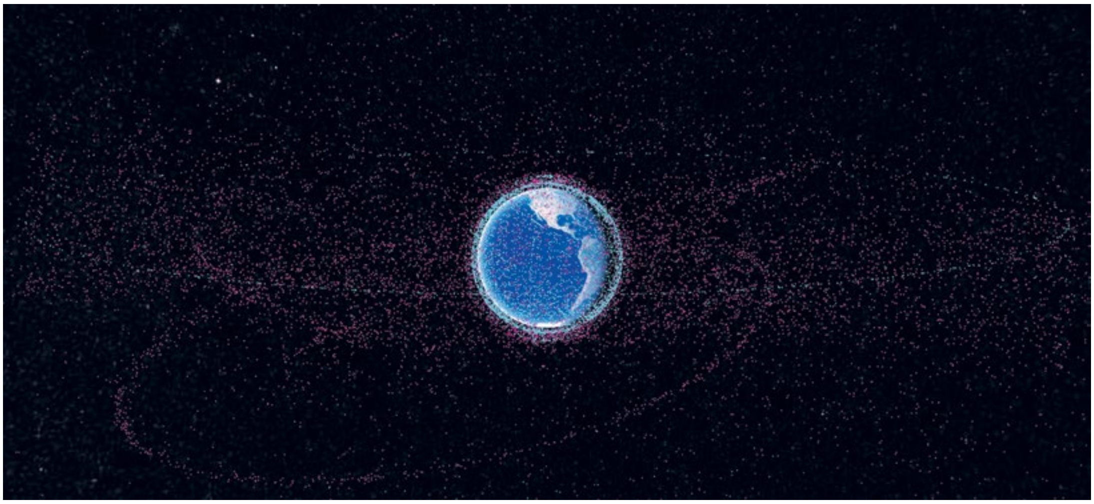
Policy, Legal, and Regulatory Issues

SPACE GOVERNANCE

Space governance has developed at the same rapid pace as the rest of the industry.[7] Within the United States, the growth of the satellite sector far outpaces the capabilities of the current licensing process. The system relies heavily on the Federal Aviation Administration (FAA) for licensing the operation of launch and reentry vehicles, and for the use of launch sites, and on the Federal Communications Commission (FCC) for communications. The demand for space-based communications is growing rapidly. In 2020 and 2021, the FCC reviewed license applications for over 64,000 new satellites, compared to a total of about 8,400 in-orbit satellites today.

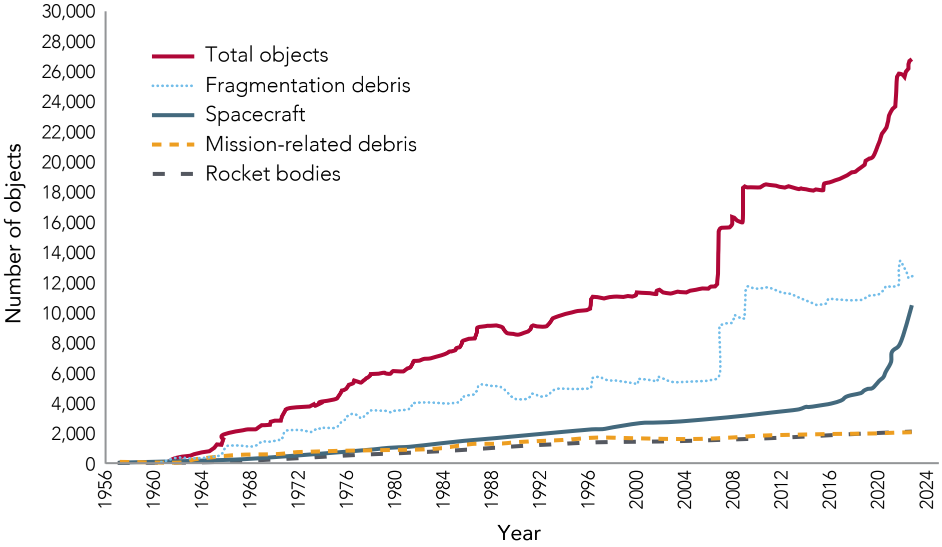
Despite their importance, space assets today are not designated by the United States as critical

FIGURE 9.5 A cloud of objects in space



Source: Privateer

FIGURE 9.6 Objects in space over time



Source: *Orbital Debris Quarterly News* 27, no. 1 (March 2023): 12.

infrastructure. Legislation has been proposed to make this designation, and as of this writing the prospects of passage are unknown.[8]

MAINTAINING SPACE ACCESS

An important grand challenge for the future of spaceflight is the preservation of space as a global resource. The near-Earth space environment is increasingly crowded, driven by lower launch costs and satellite miniaturization.

Figure 9.5 depicts objects resident in space, which include functioning satellites but also nonfunctioning satellites, or pieces of spacecraft after breakup, as well as launch vehicle components.[9] With so many objects in space, the risk of collision between two such objects is growing.

Collision has two consequences. First, if one of the two colliding objects is a useful spacecraft, it is likely to be destroyed or seriously damaged, and the likely impact velocities are so high that it is not feasible to armor satellites adequately. Second, a collision between two objects is highly likely to produce a cloud of thousands of smaller debris objects, each of which will remain in orbit to threaten still other useful spacecraft. Although a few such collisions have occurred and another few have been avoided, the possibility of a chain reaction today is relatively low. But at some point in the future, as the number of satellites being placed into orbit grows rapidly, the probability of a catastrophic chain reaction known as the Kessler syndrome will also grow. If this happens, the cloud of debris orbiting Earth resulting from such a reaction will essentially prevent space access as we know it today.

The number of objects in space has grown rapidly. Figure 9.6 shows the total number of tracked objects (each larger than 10 cm) in space since 1959. Today, there are around 35,000 such objects, of which 8,400 are working satellites—4,500 alone belonging to the Starlink satellite network. There are

2,000 nonfunctional satellites and 2,000 discarded rocket stages, and the remainder are unidentified objects.[10] There are an estimated one million fragments, between 1 and 10 cm in size.

In addition, increasing volumes of space traffic may lead to communications interference. Coordination of space activities such as orbit planning will be increasingly difficult to manage with the increase of space actors—more nations and private companies. Large satellite constellations may fill up useful orbits in ways that prevent others from using those orbits.

To reduce the impact of such factors, activities are underway that will focus on:

- **Removal of debris from orbit**, which will reduce the likelihood of collision. Requirements on actors that launch to de-orbit spacecraft shortly after they reach the end of their useful lives and active de-orbiting measures for objects such as existing discarded rocket stages will both be necessary. Some such regulatory requirements are in place, but because compliance incurs additional expenses, enforcement of these regulations is rare.

- **Automated collision avoidance systems** that will enable spacecraft to maneuver to avoid impact with space debris and other resident space objects.

- **Increased registration of launched objects** with the United Nations Register of Objects Launched into Outer Space—a registry that has existed since 1962—and bilateral or multilateral data sharing on objects to be placed in orbit would facilitate object tracking.

- **Management of space traffic**, which will require improving adherence to existing guidelines for space sustainability and strengthening international cooperation.

GEOPOLITICS, NATIONAL SECURITY CONCERNS AND CONFLICT IN SPACE

International disputes and tensions threaten the peaceful operation of satellites, space stations, and other space activities. The Outer Space Treaty was signed in 1967, at a time when the potential for the exploitation of space resources for both civilian and military purposes was not nearly as apparent as it is today. It can therefore be expected that the treaty will come under increasing pressure due to the national interests of the treaty's signatories.

The proliferation of antisatellite weapons is a major concern. To date, four nations have tested weapons capable of destroying or interfering with satellites in space—China, Russia, India, and the United States. Nations deploy antisatellite capabilities because the space capabilities of adversaries, left unchecked, provide those adversaries with military advantages. These nations can be expected to take measures to defend their own space assets while trying to degrade and deny the space assets of adversaries.

Additionally, the threat environment of the 1967 Outer Space Treaty did not anticipate cyberattacks on space missions, which can lead to data corruption, jamming, and hijacking of space intelligence providers and customers.[11] GNSS MEO-based services are especially vulnerable due to their weak signals and underline the importance of LEO-based services, as well as the exploitation of signals of opportunity—such as from mega-constellations—and passive means toward position, navigation, and timing. US Space Policy Directive-5 addresses some space cybersecurity concerns for private-sector space actors but is widely regarded as an unfunded mandate that simply adds to costs of space access. A continuing lack of governance and agreed-upon international policy thus raise the possibility of direct conflict in space.

NOTES

1. Morgan Stanley, "5 Key Themes in the New Space Economy," May 19, 2022, https://www.morganstanley.com/ideas/space-economy-investment-themes.

2. Euroconsult News, "Value of Space Economy Reaches $464 Billion in 2022 despite New Unforeseen Investment Concerns," January 9, 2023, https://www.euroconsult-ec.com/press-release/value-of-space-economy-reaches-424-billion-in-2022-despite-new-unforeseen-investment-concerns-2.

3. Alexandra Witze, "2022 Was a Record Year for Space Launches," *Nature* 613, no. 7944 (January 11, 2023): 426, https://doi.org/10.1038/d41586-023-00048-7.

4. Stevan M. Spremo, Alan R. Crocker, and Tina L. Panontin, *Small Spacecraft Overview*, NASA Technical Reports Server, May 15, 2017, https://ntrs.nasa.gov/citations/20190031730.

5. 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.

6. White House, "White House Office of Science and Technology Policy Unveils National In-Space Servicing, Assembly, and Manufacturing (ISAM) Implementation Plan," Press Releases, 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.

7. Sophie Goguichvili et al., "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.

8. Sarah Fortinsky, "Bipartisan Bill Designates Space as Critical Infrastructure Sector," *The Hill*, July 27, 2023, https://thehill.com/homenews/space/4123413-bipartisan-bill-designates-space-as-critical-infrastructure-sector.

9. *National Geographic*, "Orbital Objects," accessed August 30, 2023, https://www.nationalgeographic.com/science/article/orbital-objects.

10. Space Safety, "Space Debris by the Numbers," European Space Agency, last modified August 11, 2023, https://www.esa.int/Space_Safety/Space_Debris/Space_debris_by_the_numbers.

11. Northern Sky Research, *Global Space Economy*, 3rd ed., January 2023, https://www.nsr.com/?research=global-space-economy-3rd-edition.

10

SUSTAINABLE ENERGY TECHNOLOGIES

KEY TAKEAWAYS

- The most significant challenge to achieving sustainable energy is scale. Countries will need to source, manufacture, and deploy massive generation, transmission, and storage capabilities to meet global energy needs.
- Because global energy needs are vast, no single technology or breakthrough will be enough.
- Over-the-horizon challenges include decentralizing and modernizing the country's electricity grids and achieving greater national consensus about energy goals to enable strategic and effective R&D programs and funding.

Overview

Energy is a key strategic resource for nations. For the past few hundred years, fossil fuels have been the primary source of human energy consumption. However, humanity's reliance on fossil fuels has released enough carbon dioxide and other greenhouse gases into the atmosphere to raise the specter of significant changes in climate around the planet over the next century.

A critical aspect of minimizing the harm from increases in greenhouse gases, especially CO_2, is a transition to more sustainable energy sources over the next couple of decades. Figure 10.1 depicts the percentage of greenhouse gas emissions associated with each sector of the economy.

One critical aspect of reducing the harm from increases in atmospheric CO_2 is developing more sustainable

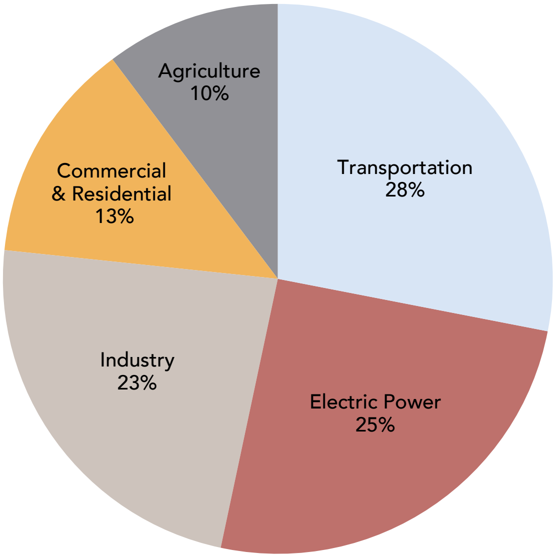


FIGURE 10.1 Greenhouse gases emitted, by economic sector

Source: US Environmental Protection Agency

energy sources such as solar, wind, hydropower, and nuclear energy.

The most significant challenge affecting the energy transition is scale. Global energy needs are large, growing, and hard to fathom. Numbers such as a billion cars, hundreds of millions of trucks, billions of tons of material to be stored or produced, billions of people to feed, and tens of millions of airplane flights per year characterize the scale of the problem. The excess CO_2 in the atmosphere is similarly characterized by large numbers: tens of billions of tons of CO_2 are produced from burning fossil fuels every year. Figure 10.2 demonstrates how the scale challenge manifests in industries such as transportation, construction, and heavy industry.

The scale challenge has two major implications. First, *no single technology or breakthrough can possibly meet the world's demands for energy*. Success will require a combination of approaches that bridge present sources, consumption, and infrastructure to a more sustainable future. Further, energy technologies must be deployed over the planet on a scale commensurate with the number of people who will use that energy. Second, *the imperative to deliver energy at scale unavoidably places an emphasis on cost*. High-cost technologies, whether new or old, and no matter how promising, cannot and will not be deployed on a wide scale.

The federal government is actively involved in investing in research and development across energy technologies. In the US government, the Department of Energy (DOE) supports innovation in energy technologies.[1] Within the DOE, the Advanced Research Projects Agency-Energy (ARPA-E) introduced entirely new methods for soliciting, evaluating, and rapidly funding new ventures, while also overseeing project progress for disruptive clean energy projects. The DOE also has specific research programs to support the development of advanced energy storage and transmission technologies, examples of which include the Energy Storage Research, Development, and Demonstration Program and the Energy Storage Grand Challenge,[2] and the Transmission Reliability Program and Smart Grid Grants,[3] respectively.

As for the economic impact of sustainable energy, one analysis indicates that doubling the share of renewable energy as a fraction of the world's energy consumption by 2030 would increase global GDP by up to 1.1 percent, or $1.3 trillion. These positive effects would be driven mostly by increased investment in renewable energy deployment, which triggers ripple effects throughout the economy.[4] Such a transition would also create 24 million jobs globally for people working in the renewable energy sector.

Domestically, eliminating air pollution emissions from energy-related activities in the United States using renewable energy sources would prevent more than fifty thousand premature deaths each year and provide more than $600 billion in benefits each year from avoided illness and death.[5]

FIGURE 10.2 The scale challenge in global energy transition

Cars

- There are an estimated 1.45 billion cars on the road today.
- Globally, one in every seven cars bought in 2022 was an electric vehicle.[a]
- Global lithium production will have to increase by a factor of 2.5 to 5 to meet expected demand for electric vehicles by 2030.[b]

Trucks

- Globally, there are approximately 217 million freight vehicles (including light commercial vehicles, medium- and heavy-duty trucks, and buses).[c]
- In 2022, 1.2% of trucks sold worldwide were electric (60,000 units).[d]

New Construction

- From 2020 to 2060, we expect to add about 2.6 trillion square feet of new floor area to the global building stock—the equivalent of adding an entire New York City to the world, every month, for 40 years.[e]
- After water, concrete is the second most-consumed material, with 30 billion tons poured each year.[f]
- If the cement industry were a country, it would be the third largest carbon dioxide emitter, following only China and the United States.[g]

Industrial Energy Use

- Heavy industry—including steel, cement, and chemical production—accounts for nearly 40% of global carbon dioxide emissions. These emissions are the hardest to decarbonize and would require both an entire change in process and building new processing plants that would require even more steel and cement.[h] Currently there are zero cement plants and only one steel plant that don't produce carbon dioxide.[i]

a. https://www.iea.org/energy-system/transport/electric-vehicles.

b. https://www.popularmechanics.com/science/energy/a42417327/lithium-supply-batteries-electric-vehicles/; https://www.sustainabilitybynumbers.com/p/lithium-electric-vehicles. Research conducted by Oxford University data scientist Hannah Ritchie.

c. https://www.shell.com/energy-and-innovation/the-energy-future/decarbonising-road-freight.html.

d. https://www.iea.org/reports/global-ev-outlook-2023/trends-in-electric-heavy-duty-vehicles.

e. https://www.iea.org/data-and-statistics/charts/global-buildings-sector-co2-emissions-and-floor-area-in-the-net-zero-scenario-2020-2050.

f. https://www.climateworks.org/blog/why-you-should-care-about-cement-and-concrete.

g. https://www.theguardian.com/cities/2019/feb/25/concrete-the-most-destructive-material-on-earth.

h. https://www.brookings.edu/wp-content/uploads/2021/06/FP_20210623_industrial_gross_v2.pdf.

i. https://www.gatesnotes.com/2022-State-of-the-Energy-Transition.

Key Developments

Substantial progress has been made in the development of several sustainable energy technologies:

- **Clean electricity generation produced by solar, wind, nuclear, and hydropower sources** For example, the cost of wind-generated electricity is substantially lower than that of fossil fuels.[6]
- **Long-distance transmission lines** Ultra-high voltage DC transmission lines are about twice as efficient at transmitting electrical power over long distances than AC transmission lines, which account for the majority of transmission lines.
- **Lighting based on light-emitting diodes (LEDs)** In contrast to incandescent lighting, LED lighting is up to ten times more efficient in converting electricity to usable light.

- **Electric car battery improvements** Electric cars today are made possible by the adoption of lithium-ion battery storage for electrical energy. Compared to the nickel-hydride batteries common in the first generation of hybrid vehicles, Li-ion batteries have about two or three times the energy density (they store more energy pound for pound), hold their charge around three times longer, and have a useful life that is about twice as long, making them more suitable for use in vehicles.

The contribution of these relatively mature technologies in the energy transition is as much a matter of public policy as scientific or technical ones. What follows are areas where the technologies themselves are not as mature.

Over the Horizon

New Technologies

LONG-DURATION ENERGY STORAGE

Sustainable energy sources such as solar and wind are intermittent. Without long-duration energy storage, the electric grid is perhaps only 50 to 60 percent sustainable. Beyond that, storage is needed and a variety of technological concepts are being researched:

- **Gravity storage** Power generated in excess of demand can be used to pump water from lower to higher levels and recovered by letting the water flow back down through generators. Large multi-ton weights can be lifted hundreds of meters and then allowed to fall gently to recover energy.

- **Thermal storage** This approach stores excess power in the form of heat, such as heating a large volume of salts to a very high temperature. When needed, that heat can be released to generate power.

- **Low-cost battery storage** Batteries beyond lithium-ion batteries are being developed, such as redox flow, Ni-H_2 gas, and Zn-MnO_2 chemistries.

None of these technologies is a silver bullet for energy storage—if economically feasible, each will fill its own niche applications. The chief challenges of all these forms of long-duration energy storage are scalability and cost reduction.

LOW-COST, HIGH-ENERGY DENSITY BATTERIES

Batteries can capture electrical energy and release it on demand. The parts of a battery responsible for capturing and releasing energy are the cathode made of one substance, the anode made of a different substance, and the electrolyte. To capture electrical energy, two different chemical reactions occur at the cathode and anode, each reacting with the electrolyte. The result is that ions released in the chemical reaction travel through the electrolyte and end up on the anode. When the battery releases energy (i.e., when it discharges), the chemical reactions are reversed at the anode and the cathode, with electrons flowing out of the anode into the light or motor the battery is powering and then back into the cathode.

Battery science is characterized by identifying better materials for the cathode, anode, and electrolyte. "Better" materials can be defined as having different physical or chemical properties (e.g., they are able to store more energy per kilogram, have lower costs, and are more available from sources friendly to the United States).

For example, the alkaline batteries used in a flashlight use different materials than the lead-acid battery used in a car. The nickel-metal hydride or lithium-ion battery in a hybrid or all-electric car is different from either of these, using still other materials.

The most significant challenge affecting all aspects of the energy transition in every sector is scale.

Batteries may be a useful way for homeowners to store excess solar power for later use, but at present they are too expensive and difficult to maintain for use as energy storage on a grid scale. For example, consider the problem of storing the world's electricity consumption for seventy-two hours; this problem sets the scale of the storage problem for sustainable electricity generation.[7] Around 200,000 gigawatt-hours of battery storage would be needed and, at about 200 watt-hours per kilogram, would require about a billion tons of battery. Consider the average US household, which uses about 900 kilowatt-hours of electricity per month; 1 kilowatt-hour is the amount of electricity that a 100-watt incandescent lightbulb uses in ten continuous hours of being on.

Lithium-ion batteries are the best batteries available today for large-scale production. The total actual and planned world production of Li-ion batteries today would be able to store 1,000 gigawatt-hours per year.[8] So, building a significant amount of battery storage for renewable energy will take several decades or even a couple of centuries at current and planned battery production capacities. In addition, lithium-ion battery costs are currently around $100 per kWh of storage capacity, while the seasonal long-duration storage that will be required must be one-tenth as costly.

The conclusion is that lithium-ion batteries will not satisfy all our needs in long-duration energy storage. Such batteries need to be able to endure tens of thousands of capture-and-release cycles, retain charge over several tens of hours, and be made of inexpensive materials. Aqueous battery chemistries, such as manganese-hydrogen batteries, for long-duration energy storage are more promising from a cost perspective—key materials such as manganese are one-tenth the cost of nickel—and they have lower life-cycle costs due to reduced maintenance needs.[9]

RENEWABLE FUELS: COMBUSTIBLE HYDROCARBONS AND BIODIESEL

Research on renewable fuels aims to create fuels that do not rely on extraction from the earth and whose burning does not release the carbon previously stored underground. Renewable fuels include combustible hydrocarbons such as biodiesel, which can be produced from animal fats or vegetable oils, and bioethanol produced from corn or algae.

Hydrogen is an important aspect of transitioning to renewable fuels. It can be directly burned without releasing CO_2, and its energy density is three times that of fossil fuels. However, for most transportation applications, frequent refueling is impractical, and to carry enough hydrogen, it must be in the form of a liquid or a highly compressed gas. In these forms, the energy density of hydrogen is significantly lower—by a factor of four—than for hydrocarbon fuels, which means that a hydrogen tank for a car needs to be four times larger to provide comparable range.

Research efforts for hydrogen storage are therefore vital if hydrogen is to play a meaningful role in the energy transition. These efforts focus on developing cost-effective hydrogen storage technologies with improved energy density that do not depend on liquefication or compression.[10] It may also be possible to use captured CO_2 combined with sustainably produced hydrogen to produce renewable hydrocarbon fuels.

Cost-effective means of producing hydrogen and carrying it with acceptable leakage from production facilities to users pose additional challenges. Currently, hydrogen is sourced from fossil fuels through processes such as naphtha reforming, natural gas steam reforming, and coal gasification. This conventional hydrogen, named gray hydrogen, has a significant carbon footprint and is not sustainable. Blue hydrogen created from methane and green hydrogen, which uses renewable electricity to generate hydrogen from water, are gaining attention.

CARBON CAPTURE AND REMOVAL

Energy is necessary for economic prosperity, and fossil fuels have been the primary source of energy for societal consumption for many decades. Fossil fuels have many advantages over other sources of energy.[11] The earliest energy sources were based on the consumption of biomass—essentially, plant material. Wood from trees in the forest was burned, for example. Animals and people performed physical labor but had to be fed from foodstuffs that were grown (e.g., grass, grains). But biomass as an energy source is limited by photosynthetic processes that capture energy in real time from the sun and is generally insufficient to support urban life, which is more population dense than most farmland.

Fossil fuels have physical characteristics that make them far superior to biomass as an energy source. Fossil fuels essentially store solar energy captured eons ago in concentrated form and carry a significantly larger amount of energy per kilogram. Such energy can be released on demand and in liquid form is especially useful and convenient in road or airborne vehicles. Coal and natural gas continue to provide a large fraction of the world's electricity.

Despite their advantages, however, humanity's reliance on fossil fuels has released enough carbon dioxide and other greenhouse gases into the atmosphere to raise the specter of significant changes in climate around the planet over the next century. Thus, a critical aspect of minimizing the harm from increases in atmospheric CO_2 is a transition to more sustainable energy sources over the next couple of decades. During this period, the United States has opportunities to consolidate its leadership in new energy technologies, strengthen national security, renew potential for economic growth, and improve equity.

Emission-free energy production will take decades to accomplish, and fossil fuels will be an appreciable (though declining) fraction of society's mix of energy sources for some time to come. In the meantime, carbon capture technology is advancing. This is the capture of CO_2 as it is being produced by the burning of fossil fuels or in industrial processes so that "new" CO_2 enters the atmosphere at a lower rate. Carbon capture technology combined with fossil-fuel-burning power plants is a way to obtain some of the benefits of fossil fuels while incurring reduced costs of CO_2 emission. By contrast, carbon removal refers to the removal of "old" CO_2 from the atmosphere. In both cases, CO_2 that is captured or removed must be sequestered in some storage facility for many decades if it is not to affect climate on Earth.

Carbon capture usually takes place at the source of emissions, such as the smokestack of a fossil-fuel-burning power plant. Source capture takes advantage of the fact that CO_2 emissions are much more concentrated at the source; once dispersed by the wind into the atmosphere, they become much harder to capture. Technologies to capture CO_2 at the source include liquid and solid materials that hold on to CO_2 in large amounts and then are sequestered and membranes that can separate CO_2 from other gases.[12] Research challenges for source capture include developing inexpensive materials for capturing CO_2 rather than other gases.

These materials are easy to handle and manage and require little energy in the regeneration process of releasing captured CO_2 for recovery. Membrane development challenges include production cost and stability requirements as well as greater permeability and selectivity for CO_2.

Carbon removal calls for capturing CO_2 directly from the atmosphere—also known as direct air capture (DAC)—at concentrations much lower than at the smokestack for carbon capture. This generally means that DAC (at least engineered DAC) uses significantly more energy in capturing a ton of CO_2 than capturing it from a point source. Engineered technologies for both DAC and point source capture rely on absorptive/adsorptive materials, though DAC materials must be optimized for use in low-concentration environments. Potentially scalable DAC approaches include biomass storage; mineralization of CO_2 using silicates; ocean alkalinization; and algae. The research challenges for DAC are scalability and cost reduction.

DIGITIZATION AND ENERGY SYSTEM INTEGRATION

In large part the result of using technologies from the early twentieth century, today's electric grid in the United States is highly centralized and operates as a single unit through the real-time coordination of power plants spread across many states. Because the ability to store electricity is minimal, such coordination must from moment to moment balance supply with demand, creating the potential for significant instability in the event of outages that would otherwise be highly localized.

The electric grid of the future will be far more decentralized and heterogeneous than the one of today. Sources of electricity will be more varied and geographically distributed as local power generation increases. Consumers of electricity will become more numerous as electrically operated systems displace systems powered by fossil fuels. Energy storage—virtually nonexistent in today's grid—will have to be managed as well. Demands for additional power will increase, requiring more power-generating facilities as well as more efficient use of existing power sources. Those demands will have to be better synchronized with timelines for generation and release of electricity from storage in ways that minimize CO_2 production.

Addressing all these challenges securely is the goal of what is generally known as the "smart grid," which will coordinate all these moving parts to increase efficiency, reliability, and resilience against attack or natural disaster.

NUCLEAR POWER

Energy can be released from the nucleus of atoms through fission, the splitting of the nuclei of certain heavy elements into components, and fusion, the merging to the nuclei of certain light elements into one nucleus. Nuclear power is an important source of emission-free electricity and has the potential to be of even greater importance in the energy transition. A deeper discussion of nuclear power is contained in chapter 6.

Challenges of Innovation and Implementation

MANUFACTURING

With many corporate and government net-zero targets established for 2040 or 2050, the next decade or so is a watershed. The United States has unparalleled capacity for fundamental research in energy. But in the equally important capacity of manufacturing at scale, the United States is no longer the world leader; China and other countries with lower operating costs control most of the manufacturing, supply chain, and critical minerals for battery and solar cell production. Since these technologies will be directly tied to the energy security of the United States, promoting domestic production will be important.

UNIVERSITY-INDUSTRY PARTNERSHIP

Universities are often the source of new innovations, but as noted earlier, many of the fundamental issues of energy are related to scale, and universities do not have the resources to effect large-scale deployments. Large energy companies, including some that have previously built their businesses around fossil fuels, are increasingly involved in the sustainable energy ecosystem. Start-ups are involved in the commercialization of research that emerges from academia. Both large and small companies have entered partnerships with academic institutions such as Stanford.

ECONOMICS AND EMPLOYMENT

Energy and economics are intrinsically linked, with the cost of energy production, energy prices, and efficiency directly influencing the economy's health and competitiveness. The energy sector is a significant part of the economy, and transitions in energy policy or sources, such as the shift to sustainable energy, can create economic winners and losers in the short term. For example, many well-paying jobs in the fossil-fuel industry are at stake in the transition to a sustainable energy regime. Economics, culture, and values are all implicated in such transitions. Manufacturing skills—essential for technology scalability—are also on the decline in the US workforce.

ENVIRONMENTAL IMPACTS

Many energy production methods, including those for sustainable energy, produce waste products that can be harmful to human health or the broader ecosystem. For example, all facilities that produce renewable energy have finite lifetimes, and at some point they need to be replaced. If these sources of energy are widely deployed—as they must be to have a meaningful impact on reducing emissions from production—they will also generate large amounts of environmental waste in the form of old windmill blades, dead solar cells, and so on. Production of biofuels is often accompanied by streams of contaminants that must be removed from the biofuels before they are shipped to consumers. The magnitude of such burdens, as well as who is responsible for them, is an important public policy concern.

Also in the category of environmental impact is the fact that many forms of sustainable energy require new acquisitions of land to build generating stations and storage facilities. For example, wind energy requires the construction of many wind turbines on large tracts of land. Residents may support windmills in principle, but "not in my backyard."

SUSTAINED FUNDING THROUGH THE VALLEY OF DEATH

The *valley of death* refers to the period after research has demonstrated the engineering feasibility of a particular innovation (a step beyond scientific feasibility) but before the innovation achieves adoption on a scale large enough to establish the viability of a business model using that innovation.

Venture capital firms are often willing to fund promising start-up companies that have generally worked out the technical bugs in their business and production processes. But in some fields, the gap between development efforts to generate prototypes to large-scale market viability is wide, and pilot projects may be necessary as an intermediate step between academic R&D. Such projects are meant to shake out technical problems that may occur only at scales significantly larger than those typically associated with prototype development. Venture capitalists are generally unwilling to invest at the larger scales that pilot projects entail, which may leave an important bridge to commercialization uncovered by external funding.

These considerations are particularly important for energy, where the importance of scale is paramount.

Policy, Legal, and Regulatory Issues

LACK OF DOMESTIC CONSENSUS ON THE NEED FOR EMISSIONS-FREE ENERGY

The federal government plays an important and large role in funding energy research and development. However, energy research requires sustained support with a long-term vision. Commercial technologies such as solar cells and batteries stem from fundamental research originating decades ago, with these technologies only now reaching fruition. For many of these technologies, large fluctuations and inconsistent support are damaging research enterprises that depend on the ability to retain knowledgeable and experienced scientists and engineers to do the relevant work.

OVERLAPPING JURISDICTIONS

Energy policy in the United States is governed by a variety of overlapping jurisdictions, such as federal, state, and local governments that have some but not necessarily final authority over the implementation of new innovations. For example, the California Public Utility Commission exercised regulatory authority in December 2022 to significantly reduce the rates at which electric utility companies are required to buy excess power that might be generated through rooftop solar cells—a move that many analysts believe would discourage homeowners from installing them in the future.[13]

NOTES

1. US Department of Energy, "Innovation," accessed October 9, 2023, https://www.energy.gov/innovation.

2. US Department of Energy, "Energy Storage RD&D," Office of Electricity, accessed September 28, 2023, https://www.energy.gov/oe/energy-storage-rdd; US Department of Energy, "Energy Storage Grand Challenge," accessed September 28, 2023, https://www.energy.gov/energy-storage-grand-challenge.

3. US Department of Energy, "Transmission Reliability," Office of Electricity, accessed September 28, 2023, https://www.energy.gov/oe/transmission-reliability; US Department of Energy, "Smart Grid Grants," Grid Deployment Office, accessed September 28, 2023, https://www.energy.gov/gdo/smart-grid-grants.

4. Rabia Ferroukhi et al., "Renewable Energy Benefits: Measuring the Economics," International Renewable Energy Agency, 2016, https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2016/IRENA_Measuring-the-Economics_2016.pdf.

5. Nicholas A. Mailloux et al., "Nationwide and Regional PM2.5-Related Air Quality Health Benefits from the Removal of Energy-Related Emissions in the United States," *Geohealth* 6, no. 5 (May 2022), https://doi.org/10.1029/2022GH000603.

6. Max Roser, "Why Did Renewables Become So Cheap So Fast?," Our World in Data, December 1, 2020, https://ourworldindata.org/cheap-renewables-growth.

7. In reality, not all sources of electricity even in a sustainable energy world will be intermittent, but many will be. The required storage may be only forty-eight hours rather than seventy-two, and so on. But these figures set the required scale.

8. Steve Hanley, "CATL M3P Battery Production Begins, DOE Predicts 1000 GWh of North America-Built Batteries by 2030," CleanTechnica, March 26, 2023, https://cleantechnica.com/2023/03/26/catl-m3p-battery-production-begins-doe-predicts-1000-gwh-of-us-built-batteries-by-2023.

9. Wei Chen et al., "A Manganese–Hydrogen Battery with Potential for Grid-Scale Energy Storage," *Nature Energy* 3 (April 2018): 428–35, https://doi.org/10.1038/s41560-018-0147-7.

10. "Hydrogen Storage," US Department of Energy, accessed August 30, 2023, https://www.energy.gov/eere/fuelcells/articles/hydrogen-storage-fact-sheet.

11. Samantha Gross, "Why Are Fossil Fuels So Hard to Quit?," Brookings Institution, June 2020, https://www.brookings.edu/articles/why-are-fossil-fuels-so-hard-to-quit.

12. International Energy Agency, "About CCUS," April 2021, https://www.iea.org/reports/about-ccus.

13. Julie Cart, "California's Residential Solar Rules Overhauled after Highly Charged Debate," CalMatters, December 15, 2022, https://calmatters.org/environment/2022/12/california-solar-rules-overhauled.

11

CROSSCUTTING THEMES AND COMMONALITIES

Chapters 1 through 10 addressed ten individual technology areas. This chapter pulls together some common themes that cut across them. Of course, the scientific issues at play are different because the science is different. However, there are important similarities in how people and institutions make progress that are often lost when each field is considered in isolation.

The Value and Risk of Technological Progress

Takeaway *Innovation that emerges too fast threatens the legitimate interests of those who might be negatively affected by such innovation, while innovation that moves too slowly increases the likelihood that a nation will lose first-mover advantages.*

New technologies typically bring two types of benefits. First, they can enhance or improve existing processes. Second, they can enable entirely new functions—addressing problems in novel ways and solving challenges that people did not even know they had. In the first half of the twentieth century, for example, polio afflicted thousands of people worldwide. The iron lung was invented in the 1930s to help polio victims breathe, and over the next twenty-five years, improvements were made to the iron lung.[1] But the groundbreaking Salk vaccine in 1955 brought an entirely different way to defend people against polio. Within a few years, use of the iron lung dropped to nearly zero.

Manufacturing provides another example. For decades, large-scale manufacturing has relied on the idea of an assembly line to fabricate essentially identical models of the same product. Workers were originally all human. Then robots began replacing them, performing many assembly-line tasks more rapidly and accurately while reducing production costs. In the past two decades, a complementary fabrication paradigm has emerged: custom, on-demand manufacturing of products in small quantities using 3-D printing, or what's known as additive manufacturing. This new paradigm enables production that is far more localized and customized, though it does not yield the economies of scale that mass production offers.

Technological progress also brings risks—risks of moving too fast or too slowly. Innovation that emerges too fast threatens to disrupt the often-delicate balance that has been established among many national, organizational, and personal interests. As we are seeing today with AI, the rush to deploy new capabilities may give short shrift to issues such as safety, security, employment, values, ethics, societal impact, and geopolitics. On the other hand, 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. In both cases, policy measures are often needed to steer outcomes in a more optimal direction.

The road from scientific discovery to useful application is often rockier than expected as well, with would-be innovators finding that the realization of the benefits promised to investors and customers actually entail greater costs, deliver fewer capabilities, and take more time than anticipated. Furthermore, it may well be that only upon delivery of new products do other risks become apparent, with innovators facing issues of ethics and equity, privacy, and increased challenges to health, safety, and security—all risks that could lead to an erosion of trust in their services or capabilities.

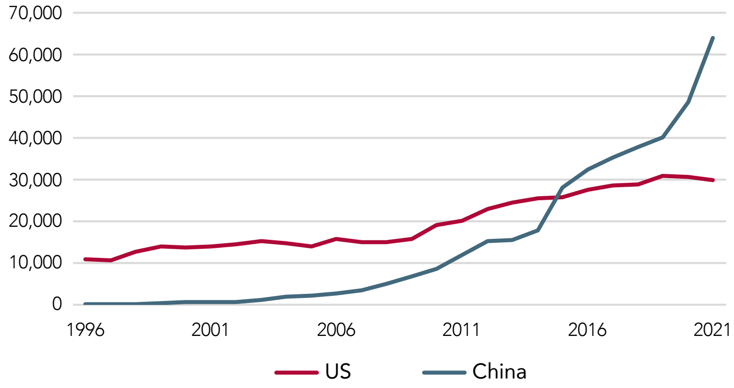
The Central Importance of Ideas and Human Talent in Science and Technology

Takeaway *Human talent plays a central role in generating the ideas for innovation, it can be found all over the world, and it cannot be manufactured at will.*

From time to time, lone scientists working on their own achieve breakthroughs on very difficult problems. But it is far more common that successful science and technology (S&T) innovations are a result of a well-functioning collaborative effort that can bring to bear a broad range of cognitive styles and disciplinary expertise.[2]

Scientific progress obviously benefits from new ideas. New ideas are created every day by talented Americans, but Americans do not have a monopoly

FIGURE 11.1 Number of patents per year



Source: World Intellectual Property Organization, "WIPO IP Statistics Data Center."

on the creation of new ideas. As one metric, consider that China's production of patents reached parity (100 percent of US patent production) around 2015 and (as of 2018) surpasses that of the United States (see figure 11.1). Even allowing for the possibility that historically, Chinese patents might be of generally lower quality than those of the United States, the trend line is clear.

Other nations are also investing heavily in research and development (R&D), which is a critical source of new ideas. According to the National Science Board, most of the world's R&D expenditures occur in a few countries, with the United States accounting for 27 percent of global R&D in 2019, followed by China (22 percent), Japan (7 percent), Germany (6 percent), and South Korea (4 percent).[3] At the same time, the concentration of R&D expenditures continues to shift from the United States and Europe to countries in East, Southeast, and South Asia. This trend is consistent with the observation that an increasing share of the world's patents are shifting over to Asia—particularly to China.[4]

How can the United States take advantage of ideas produced in other countries? One obvious way is to read the scientific and technical literature produced by scientists abroad, and that does happen in abundance. But it is well known that people are a much more effective information transfer mechanism than papers.

For example, informal interviews with Stanford faculty across most of the technology areas yielded two important points regarding the value of direct, in-person interactions with foreign scientific colleagues. First, these interactions enable them to learn things they could not learn simply from reading papers published by the same people, as papers often do not capture vital "tacit knowledge" that enables researchers to build upon the work of others. Second, they are able to develop a much better understanding of the scope and nature of progress made and not made by their foreign colleagues—an understanding that would not result simply from reading the literature.

In fact, the same point—that ideas move much more effectively when conveyed through people than through papers—underscores how the United States actively benefits from foreign scientific input. Skilled immigrants support American innovation today. For example, immigrant college graduates receive patents at double the rate of native-born Americans.[5] In part, this is explained by a higher proportion of immigrant students pursuing STEM education in the United States. Technology companies in the United States also rely on immigration visas to bring foreign scientists and engineers to work in the United States.

Conversely, the United States suffers when foreign scientific inputs are curtailed. For instance, a Harvard Business School study found that pro-migration changes to immigration policy significantly increase innovation within a country—as measured by the production of patents—while changes that discourage immigration lead to significant declines in patent production.[6]

Human talent capable of creating ideas in S&T cannot be manufactured at will. It must be domestically nurtured or otherwise imported from abroad. Today, both paths to growing the requisite talent base to sustain and grow US innovation face serious and rising challenges. Test scores clearly show declining performance in STEM subjects in K–12 education, both in absolute terms and in comparison with other countries.[7]

Regarding US STEM education, the US Department of Defense noted in 2021 that improving the capacity and resilience of the defense industrial base requires more workers trained in the skilled trades and in STEM.[8] Yet bias against careers in the industrial trades among parents and educators has shrunk the pool of potential workers, and adverse demographic trends have led to an aging-out of a skilled workforce with irreplaceable knowledge. The Defense Department also noted the dearth of trained software engineers working on classified projects was in part because they must be US citizens.

At the same time, US immigration policies discourage or prevent foreign S&T talent from working here. Current policies are facilitating a shift of skilled immigration and associated multinational R&D investment toward other countries. Tightening immigration also can prevent companies from hiring enough skilled workers to operate their R&D facilities, increasing their incentives to relocate abroad.[9]

A significant portion of academic researchers are PhD students and professors who have immigrated to the United States to seek better educational and research opportunities. It is crucial to establish a better pathway to permanent residence upon graduation for PhD students on student visas so that the United States does not lose highly trained workers. The United States and universities invest heavily in the education of STEM graduate students, and it would be wise to find a path to allow these scientists and engineers to work and live in the country permanently. Furthermore, if ambitious goals in building up the semiconductor industry, biotechnology, or decarbonization are to be met, then increased investment is needed in the labor force. These industries hire highly trained workers who have advanced degrees. Research funding supports not only the scientific outcomes but also an essential method for training highly skilled engineers and scientists.

Finally, it is important to realize that the global talent challenge is not just about China. US allies and partners also compete for technology talent from around the world. For example, Canada has always had an immigrant-friendly policy that attracts foreign-born graduates of US universities—nearly forty thousand such individuals were recruited to Canada from 2017 to 2021.[10] More recently, Canada introduced its Tech Talent recruiting program in June 2023.[11] This program targets tech workers in the United States who hold US H-1B nonimmigrant visas, providing them more favorable terms. A 2022 survey of almost 1,500 global leaders hiring tech professionals in the United States, United Kingdom, Germany, France, the Netherlands, and Sweden found that more than one-third were searching globally.[12]

A report from Korn Ferry Consulting asserts that "human capital is the greatest value creator available to organizations,"[13] also finding that every dollar invested in human capital adds $11.39 to GDP. The same holds true for advances in science and technology more generally: the most important ingredient for S&T advances is human talent because human talent is the goose that lays the golden eggs of ideas and innovation.

None of these comments are intended to suggest that concerns about foreign appropriation of American intellectual efforts are unfounded. But the fact that American S&T efforts are deeply connected to those of the rest of the world is overall an accelerator of those efforts rather than a brake on them. Using an axe to impose blanket restrictions on engaging foreign scientists when a surgical scalpel is needed to curb only the issues that warrant serious concern is a sure way to reduce the effectiveness of US scientific efforts.

The Changing Role of Government regarding Technological Innovation

Takeaway *The US government is no longer the primary driver of technological innovation or funder of R&D.*

Many technological advancements—such as satellites and access to space, the development of jet engines, and the emergence of the semiconductor industry in Silicon Valley—have their roots in US government financial support and advocacy. But in many fields today, the US government is no longer the primary driver or funder of R&D.

Private companies have taken up much of the slack. For example, while the US government once used its own rockets to launch satellites, it now often does so by contracting with companies that provide access to space as a service. These companies, 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, for a several-year period in the mid-2010s, the United States was entirely dependent on Russia to transport American astronauts to the International Space Station. More recently, the Starlink satellite communications network has been an essential part of Ukrainian battlefield communications; however, the CEO of Starlink curtailed Ukrainian access on a number of occasions in ways that affected Ukrainian battlefield strategy.[14] Such concerns are most serious when there is only one or a small number of private-sector providers of the services in question.

Many US officials recognize the growing role of a handful of private actors behind influential innovations in technology. They believe that supporting closer public-private sector cooperation, as well as informed government regulation of emerging technologies, is a pressing imperative. Even if the government cannot rely on its own capabilities to remain at the forefront of technological innovation, it still has an important role to play in funding and promoting R&D, facilitating the broad adoption of key innovations and standards, and convening coalitions of like-minded actors both domestically and internationally.

A Trend toward Increasing Access to New Technologies Worldwide

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

Access to technologies such as synthetic biology, robotics, space, and blockchain often spread from rich nation states and large corporations to less wealthy nations, smaller corporations, universities, and individuals. Even innovations that are American born—an increasingly rare occurrence—are unlikely to remain in the exclusive control of American actors for long. Many emerging technologies exhibit a long-term trend of riding a declining cost-curve over time, making them accessible to an ever-larger set of individual actors. Export controls may delay this spread in some cases for a limited time, but the overall trend is toward decentralized access.

This trend has several implications:

- **Greater policy complexity results from more actors.** Other actors, both state and nonstate, will have capabilities to challenge US interests that they did not have before.
- **Technological advantages will diminish.** The United States may have the most technologically advanced capabilities, but even the more rudimentary instantiations of these capabilities available to other actors eliminate monopolies and narrow the relative advantages once enjoyed by the United States.
- **Winning isn't winning anymore.** The old paradigm of "winning" a technological race to achieve gains that last decades and are hard to replicate—traveling to the moon, developing the atomic bomb—will be replaced by the paradigm of constant competition.
- **More diversity in bureaucracy and ethics has consequences.** Actors less subject to bureaucratic and ethical constraints will be able to exploit technology more nimbly and adapt more rapidly to conditions on the ground.

On the other hand, for physical technologies whose effectiveness depends on deployments in large quantities, geography still plays a role. Natural resources such as rare-earth metals are geographically constrained, and production facilities for physical artifacts still matter.

It may be possible to extend periods of American monopoly on certain technologies, but not indefinitely. Such extensions can have valuable short-term benefits, not least because they buy time for US policymakers to better anticipate a world of democratization. But all too often buying time becomes an end unto itself, and actions to craft a better policy—such as improving targeted immigration reform and sustainability—are not taken.

To be sure, there are probably exceptions to this democratization trend. One is the first appearance of an emerging technology. At such a point in time, the democratization process has not yet begun, at least not in full force, and it may indeed be that the technology in question will be characterized by the dominance of a few key actors.

For example, the training of large language models (LLMs) from scratch is a new capability still only belonging to a few large companies, and it is completely out of reach even for large coalitions of top research universities. On the other hand, research is already underway to build applications based on foundational models at much lower cost. In many cases, the approach taken is to fine-tune a foundational model that has already been trained from scratch, thus building on previous efforts. Therefore, it may well be that this particular exception is apparent only because of the quirk of timing, and LLMs will also be further democratized as time goes on.

In other cases, the impact of emerging technologies depends on the scale of deployment. Universities may be able to develop sustainable energy technologies such as better batteries, but they lack the infrastructure to manufacture them at scale, an enterprise that requires enormous investments from the private sector.

Synergies between Different Technologies

Takeaway ***The synergies between different technologies are large and growing as advances in one technology often support advances in other technologies.***

The technologies described in this report span a broad range, but most have in common a synergistic relationship to other technologies. Improvements in technology A can be used to improve the performance of technology B, while improvements in B help C— then C and B together can help improve A.

For example:

- AI contributes to advances in synthetic biology in addressing the protein-folding problem, predicting protein shape from the DNA sequence of base pairs.
- AI helps to screen many candidate compounds to predict the ones most likely to exhibit desirable properties for materials science.
- 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.
- New materials are important in space research for the construction of spacecraft and satellites.
- New materials are needed to enable the development of neural probes that can send and receive electrical signals in neural tissue.
- Energy technologies help to improve the performance of robotics and spacecraft.
- Synthetic biology can build organisms that produce certain specialized materials.
- Cheaper semiconductors have driven down the cost curve of DNA sequencing, which itself is a fundamental technology for synthetic biology.[15]

Some of the S&T areas in this review—AI, synthetic biology, materials science, and energy—have a foundational flavor impacting a variety of problem or technology domains.[16] Others are better characterized as technology applications—space and robotics, for example—which focus on solving specific problems through an artful blend of a number of technologies. But these are differences of degree, not of kind, and even more foundational technologies can benefit from advances in different technology applications.

Nonlinear Paths from Research to Useful Application

Takeaway ***The traditional "linear" model of R&D, in which basic research leads to applied research, which then leads to development and prototyping, which finally leads to novel and useful products or services, is only one model for how societies obtain value for investments in technology innovation.***

The traditional "linear" model of R&D in science and technology starts with basic research leading to applied research, which then leads to development and prototyping, which finally leads to marketable products.

- Basic research is activity aimed at fundamental scientific understanding without any particular applications in mind.
- Applied research is activity to deepen this scientific understanding with an application area or specific problem in mind.
- Development is activity that builds on scientific understanding to construct engineering prototypes and proofs-of-concept.

This model of scientific development has a long history, but it is by no means the only path. Many analysts argue this model is so unrepresentative of how scientific development actually proceeds that it can be harmful to the scientific enterprise.

Other models are less linear in nature; they acknowledge and even exalt the need for feedback between the various activities. For example, some problems or application areas are so challenging that they entail obtaining a deeper fundamental scientific understanding of a real-world phenomenon—what some call "use-inspired basic research." One of the most famous examples is the work of the French chemist Louis Pasteur on milk spoilage, an applied problem that required advancements in fundamental biological science regarding bacterial processes.

In other cases, deep fundamental scientific knowledge may be necessary in technology areas adjacent to the primary problem of interest. New drugs that are effective against cancer tumors do no good if they cannot be delivered in lethal concentration to the tumor. Therefore, research on drug delivery mechanisms is as important here as the development of new anticancer compounds.

The Relationship of Political Regime Type to Technological Progress

Takeaway ***Democracies provide greater freedom for scientific exploration, while authoritarian regimes can direct sustained funding and focus to technologies they believe are most important.***

Technological innovation occurs in both democracies and autocracies, but different regime types face different advantages and challenges. True democracies enjoy the rule of law and a free flow of ideas and people, as well as the ability of individuals to pursue research goals of their own choosing. Perhaps most important, because failure in a democracy does not lead to persecution or necessarily result in professional ostracism, individuals are freer to experiment and explore.

Authoritarian regimes are more aptly characterized by the rule of the state, or the whim of a single "supreme leader." This can lead to the constrained flow of ideas, coercion to limit individual freedom of action and thought, and top-down direction to explore only topics of interest to the state. In this environment, failure may carry very high consequences for individuals. Under such circumstances, it would be understandable if individual scientists limit themselves to studying "safe topics" for which failure to make progress is unlikely.

Yet it must be noted that authoritarian regimes have the advantage of being able to direct funding and public attention to problems that they believe are important. They can sustain that focus for long periods of time more independently of short-term considerations such as profit or politics. To the extent that technology-based solutions are known, authoritarian leaders can exploit that knowledge to implement those solutions, regardless of any downsides.

Successful innovation requires both an exploration of the relevant space of possible solutions, to eliminate paths that will not lead to viable solutions, and an exploitation of viable solutions to focus resources on specific problems deemed important to the regime.[17] Competitors such as China have taken advantage of US scientific exploration in many domains through means both legal and illegal and have gone on to exploit that knowledge through a variety of commercial and military efforts.

Attempts to obtain some of the benefits of a more centralized direction to the technology policy efforts of the United States have been described as steps in the direction of adopting an industrial policy. Critics often argue that such efforts unduly interfere in a free market and that picking "winners and losers" leads to inefficiencies. Advocates argue that only through such action will the United States be able to offset some of the advantages that authoritarian nations would otherwise enjoy over it. The public policy problem is acknowledging some truth in both perspectives and seeking an appropriate balance consistent with both American values and economic competitiveness.

Punctuated Technological Progress

Takeaway *Technology often progresses in fits and starts, long periods of incremental results are followed by sudden breakthroughs, and the speed of change is hard even for leading researchers to anticipate.*

Taken as a whole, technological progress exhibits a variety of patterns. Some technologies have demonstrated consistent progress for extended periods. For example, semiconductor technology is characterized by Moore's law, an exponential reduction in the cost of semiconductors over time. Solar cells and LED efficiency have followed similar cost reduction curves.

Other technologies have demonstrated much more uneven rates of progress. These technologies see long periods of incremental development and refinement that are punctuated by short bursts of radical innovation. In some cases, the bursts are the result of some particular breakthrough—examples might be the emergence of the personal computer in the 1980s or the World Wide Web in the early 1990s. In other cases, the bursts are due to the simultaneous availability and maturity of several key technologies that are required to make significant progress in some other technological domain. Here, an example might be electric cars, where battery technology, lightweight materials, sensors, and computing power have come together to make such cars more economically feasible.

When punctuated progress characterizes a technology, forward projections of progress based on past rates may well be misleading. Successful forecasting depends on familiarity with a wide variety of technologies precisely because it is hard to predict which specific technologies will prove critical. Indeed, even experts in a given field can be surprised by the rapidity of progress, as has happened in the last year with artificial intelligence and applications such as ChatGPT and large language models. Geoffrey Hinton, a 2018 Turing Award winner for his work on artificial intelligence, said, "I have suddenly switched my views on whether these things are going to be more intelligent than us. I think they're very close to it now and they will be much more intelligent than us in the future."[18] This sentiment is shared by other experts in LLMs, who have told us they too have been stunned by the speed of advances in their own field.

While there is broad recognition that we may be at the cusp of a moment of radical technological change across a number of fields (AI, synthetic biology, nuclear fusion), the precise contours, speed, and implications of this moment are much harder to ascertain.

Nontechnological Influences on Technological Innovation

Takeaway *Technology applications in society require scientific proof-of-concept, engineering feasibility, economic viability, and societal acceptability.*

Technology plays a key role in supporting and advancing national interests, and headlines in the news often tout scientific breakthroughs that offer opportunities to solve societal problems and to improve the quality of life. But to play a valuable role, any given technology application must demonstrate not only technical feasibility but also economic viability. It must also be acceptable to the relevant constituencies, including the public at large. People and organizations must be able to adapt to its use, despite the disruption it may cause. The requirements and burdens imposed by law, policy, and regulation must be compatible with widespread adoption and use of the application.

There is often a large gap between a demonstration of scientific feasibility—let's call it Q—and a product or a service based on Q that is useful to society. Press reports of scientific breakthroughs often give the impression that useful exploitations of these breakthroughs are just around the corner. That is almost never true. Scientific feasibility is a necessary prerequisite, but it may well be that engineering or economic feasibility does not follow. After scientific proof-of-concept is achieved, engineering feasibility must be demonstrated, which includes considerations of cost and ease of use.

Scientific proof-of-concept is only the first step. Engineering feasibility must also be demonstrated, which includes considerations of cost and ease of use. Take the case of the technical success of early attempts to build supercomputers out of superconducting components that required liquid helium to cool them. Technical feasibility was demonstrated, but the liquid helium requirement would make these computers difficult to deploy and use in practice; alternative computing technologies appeared to offer comparable performance at lower cost.

In other cases, engineering feasibility can be demonstrated but cost considerations must first be resolved. For example, when carbon fibers were first being investigated in the laboratory, they cost $10 million per pound—clearly infeasible for large-scale use.[19] A substantial amount of work was required to reduce the cost by what is today a factor of a million.

In still other cases, it may prove too difficult to develop a manufacturing process to build a device based on Q, or the materials used to demonstrate Q are too expensive or rare to support large-scale production. Less expensive or cumbersome alternatives to devices based on Q may be available, thus reducing the marketplace viability of Q-based devices.

Societal acceptability matters as well. In Europe, though much less so in the United States, genetically modified organisms as food are highly controversial, and concerns over their safety have prevented the uptake of GMO foods consumed widely in the United States. The psychology of individuals and cultural practices and beliefs of a community or society also contribute to the adoption and use of any given technology application. The essential point here is that technology in society is not just about the technology.

Lastly, given that some technological demonstrations of scientific feasibility do not advance to the marketplace and become "orphaned," an important public policy question is how to manage them.

For example, a start-up company may be established to commercialize Q. If the company fails in the marketplace for economic reasons, a competitor or another nation with a different cost structure may be able to make it economically viable—and that competitor's interests may not align with those of the United States. What if the competitor is a bad actor and simply buys the now-defunct start-up, thereby acquiring the rights to the intellectual property underlying Q?

The Role of Universities in Tech Innovation

Takeaway *US universities play a pivotal role in the innovation ecosystem that is increasingly at risk.*

The US infrastructure for funding and conducting R&D (to which innovation is closely related) is broad and deep. For example, the private sector is the second-largest supporter of R&D in the United States (the first is the federal government).[20] Entities such as Bell Laboratories, IBM's Thomas J. Watson Research Center, and Xerox PARC once performed substantial amounts of basic scientific research. Today, their present-day equivalents focus most of their R&D efforts on process and product development closely related to the bottom line of their parent companies. This focus has two important implications. First, companies tend to focus their efforts on research with foreseeable commercial application, not frontier or fundamental research where the connection between breakthroughs and application may not be apparent and where it may take years, if not decades, for a technology to mature. Yet companies depend on nonindustrial research. For example, 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.[21] Second, corporate R&D outputs tend to be restricted and proprietary to preserve any market-competitive edge that they may afford to the company that paid for them.

The federal government also operates a large number of laboratories and federally funded research and development centers. For example, the Department of Energy operates seventeen national laboratories for conducting research and development that serve the department's core missions in energy, science, national security, and environmental stewardship.[22] These labs specialize in particularly difficult problems that fall beyond the capabilities of private industry or individual universities. (The fusion breakthrough described in chapter 6 was conducted at the Lawrence Livermore National Laboratory.) Nor is the Department of Energy unique in having mission-driven laboratories that actually conduct research and development work; the Department of Defense, NASA, the National Institutes of Health, and the Department of Commerce are just a few of the departments with their own mission-driven labs.

But the role of universities is unique and pivotal in the innovation ecosystem, and this role is often underappreciated. For example, in contrast to research done in mission-driven federal laboratories, the scope and breadth of research conducted at major research universities tends to be much larger than that conducted at federal labs simply because the various foci of university research are not constrained by particular missions.

In contrast to the private sector, university research is almost always open, enabling would-be innovators to take advantage of it. Open research promotes transparency, accountability, collaboration, reduced duplication, and wider impact. By sharing study details, data, and results openly, researchers allow others to verify, replicate, and build on their work more easily. This accelerates discovery as more minds can work on problems in a collaborative way, with reduced redundant efforts. Published open research also reaches more people, magnifying its educational and societal benefits.

The role of universities in building the national economy has been recognized since 1862 with the passage of the Morrill Act establishing land-grant universities. Government-supported university research made a critical difference in World War II on technologies such as radar, proximity fusing, computers—and the atomic bomb. University research has since generated knowledge whose exploitation creates new industries and jobs, spurs economic growth, and supports a high standard of living, while achieving national goals for defense, health, and energy.[23] University research has been a rich source of new ideas, particularly for the longer term, and universities are the primary source of graduates with advanced S&T skills.

Universities have the mission of pursuing high-risk research that may not pay off in commercial or societal applications for a long time, if ever.[24] For example, 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. Although this line of research was abandoned after a decade or so, it ultimately gave rise to the work in AI on deep learning several decades later. The term "mRNA vaccines" entered the public lexicon in 2021 when COVID-19 vaccines were released. Yet development of these vaccines was built on university research with a thirty-year history.[25] 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.

Finally, the increasingly blurred distinction between fundamental research and export-controlled research is creating challenges in academia in fostering international collaboration, particularly in fields such as the semiconductor industry, nanotechnology, AI, and neuroscience. Some researchers are concerned that fundamental research could now be considered export controlled and may steer clear of foreign collaborations out of an abundance of caution. This policy ambiguity can deter collaborations and create obstacles for non-US researchers wishing to contribute to work in the United States. This is particularly concerning as international cooperation could expedite progress in emerging fields like nanomaterials, where countries like Korea are making significant strides, especially in biomedical applications. These policy issues, widely recognized among the research community, underscore the urgent need for clarification and reform to advance research and promote effective international collaborations.

The Structure of Research and Development Funding and the Valley of Death

Takeaway *Sustaining American innovation requires long-term government R&D investments with clear strategies and sustained priorities, not wild swings from year to year, which is increasingly common.*

Budget is one obvious aspect of government funding for R&D. But three other aspects deserve as least as much attention. First, government has an important role in funding important research with long horizons, as industry is not generally structured to support long-term R&D efforts. Such government-funded research should generally be regarded as precompetitive in nature.

Second, wide swings in funding from year to year—increasingly common in government funding—are antithetical to a systematic R&D effort. In a free market economy, talented scientists can choose where and in what domains to work, and they have a natural aversion to work environments that do not provide stability. Therefore, wide swings in funding have the effect of driving away the scientific talent that can best find employment in that field elsewhere.

Third, the so-called valley of death remains a significant problem. This refers to the period after research has demonstrated the engineering feasibility of a particular innovation—a step beyond scientific feasibility—but before the innovation achieves adoption on a scale large enough to establish the viability of a business model based on the innovation.

When an innovation is first offered to customers, it is expected to provide new functional capability. Its cost matters as well, as the new functional capability may or may not be worth the price of adopting it. At one cost, a potential customer might choose not to acquire it, while at a lower cost, that customer may well do so.

If the initial cost is too high, customers will be scarce and the firm producing or providing the innovation is likely to fail commercially if it does not receive funding from a source not related to production to stay afloat. But it often happens that the per-unit production cost will decrease as the total cumulative volume of production increases. Known as the learning curve in manufacturing, this phenomenon is primarily due to the efficiencies and knowledge gained from repetitive production processes. Such cost reduction is particularly important when, as is true for energy production, significant societal benefits accrue and new technologies are deployed at scale.

Research funding generally disappears after feasibility has been demonstrated. For a firm then to get through the valley of death, it must either demonstrate the viability of its business model to investors who believe in the promise of the innovation or attract enough customers on its own to sustain it. True commercial viability is unlikely to start until the per-unit cost has dropped to levels affordable by most would-be customers.

While in the valley of death, it is typical that no party is willing to invest the minimum level for product or manufacturing refinement to continue, and projects often have to stop or progress much more slowly than before. In some cases, the innovation never scales beyond the initial stages, regardless of its technical sophistication or desirability.

These points suggest bridge funding may help in establishing commercial viability. The sticking point, however, is the difficulty of distinguishing between real innovations that would be truly valuable if only they could get through the valley of death and ersatz innovation look-alikes for which valley-of-death concerns are merely a smokescreen to cover up their genuine inadequacies and problems in the face of market realities.

A firm's failure to pass through the valley of death may also have competitive implications internationally. Such a firm is ripe for acquisition by foreign competitors with deeper pockets who may be willing to invest in innovative products that have not yet reached market viability. Chinese investors, for example, were successful in acquiring Atop Tech, a firm with an automated designer capable of producing high-end microchips, after it went bankrupt in 2017. This transaction failed to elicit any reaction from the Committee on Foreign Investment in the United States (CFIUS), despite its mandate to review and, if necessary, block certain transactions involving foreign investment that may impact US national security. The Foreign Investment Risk Review Modernization Act of 2018 was enacted in part to improve the ability of CFIUS to review just such transactions.

A new funding model, known as focused research organizations (FROs), seeks to fill the gap inherent in the valley of death. The FRO provides funding to assemble scientists and engineers with the required expertise to rapidly prototype and test materials and technologies for their applications. One initiative to support FROs was launched in 2021, Convergent Research, a nonprofit organization with the mission of incubating and funding new FROs. In March 2023, it received $50 million in philanthropic donations to launch two new FROs.[26]

Cybersecurity

Takeaway *Researchers working in highly competitive environments who neglect cybersecurity place their research progress at risk.*

Cybersecurity refers to technologies, processes, and policies that help to protect computer systems, networks, and the information contained therein from malicious activities undertaken by adversaries or unscrupulous competitors. It is often believed that cybersecurity is an issue that primarily affects private-sector businesses and government, but the world of academic R&D faces a variety of cyber threats as well.

One important cybersecurity interest is in ensuring the integrity of data. Scientific experiments produce data, and if important data are deleted or destroyed, scientific progress can be significantly retarded. A possibly more worrisome scenario is that the data are altered in hard-to-detect ways that subtly and invisibly skew the subsequent work based on that data, possibly putting scientific investigators on the wrong track and wasting significant effort. Adversaries or competitors seeking to delay scientific work have significant incentives to engage in activities that could compromise data in this manner.

Similar comments apply to the computer programs used to analyze data. If a computer program is maliciously altered in a subtle way, it may be a long time before the alteration is noticed. Once such an alteration is noticed, all previous analyses performed using that program are inevitably called into question.

A second cybersecurity interest is in ensuring the confidentiality of various work products, such as datasets and working papers. Datasets may have been collected under promises of confidentiality or nondisclosure agreements, and unauthorized access to such datasets quite possibly violates such promises or agreements. Premature disclosure of working papers can compromise claims of priority, an important currency in which academic R&D trades. Additionally, draft working papers are often incomplete, inconsistent, or downright wrong and are not in any sense defensible—premature disclosure of such papers as though they did in fact reflect completed work is a nightmare of any scientist.

Many laboratories rely on computers to control or supervise data collection from various instruments. Compromising these computers through a cyberattack could cripple data collection efforts or corrupt the data being collected. The instruments in question could also be damaged by hacking the controlling computers.

Technical safeguards are available for most cybersecurity problems, of which the above are just a sample. But especially in academic laboratories, maintaining and operating such safeguards consistently calls for a serious management effort to impose the necessary discipline on all those working in those labs. Such discipline often conflicts with informal laboratory cultures that stress collegiality, openness, and flexibility.

A second cyber-related threat to the R&D enterprise is selective targeting of key personnel working on important research projects. It is a matter of public record that a number of Iranian nuclear scientists have been killed since 2007, reportedly because they were associated with Iran's nuclear program.[27] But assassination is not the only form of targeting. Much less violent forms of targeting could involve what might be broadly termed harassment, which often originates in or is perpetrated through cyberspace. For example, compromising the personal life of a principal investigator (e.g., draining bank accounts, interfering with the investigator's personal finances, threatening the investigator's family) can all be accomplished through the internet. Dealing with such matters will inevitably reduce the work effectiveness of the individuals targeted in this manner. Calling into question the investigator's professional conduct and ethics is another approach that could have comparable effectiveness.

NOTES

1. Science Museum, “The Iron Lung,” October 14, 2018, https://www.sciencemuseum.org.uk/objects-and-stories/medicine/iron-lung.

2. Keith Sawyer, *Group Genius: The Creative Power of Collaboration* (New York: Basic Books, 2007).

3. Amy Burke, Abigail Okrent, and Katherine Hale, *The State of U.S. Science and Engineering 2022* (Arlington, VA: National Science Board, National Science Foundation, 2022), https://ncses.nsf.gov/pubs/nsb20221.

4. World Intellectual Property Organization, “World Intellectual Property Indicators: Filings for Patents, Trademarks, Industrial Designs Reach Record Heights in 2018,” October 16, 2019, https://www.wipo.int/pressroom/en/articles/2019/article_0012.html.

5. Jennifer Hunt and Marjolaine Gauthier-Loiselle, “How Much Does Immigration Boost Innovation?,” *American Economic Journal: Macroeconomics* 2, no. 2 (April 2010): 31–56, https://doi.org/10.1257/mac.2.2.31.

6. Dany Bahar et al., *Talent Flows and the Geography of Knowledge Production: Causal Evidence from Multinational Firms*, Harvard Business School, Technology and Operations Management Unit, Working Paper no. 22-047, last modified December 23, 2022, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4005693.

7. Amy Burke, Abigail Okrent, and Katherine Hale, “U.S. and Global STEM Education and Labor Force,” in *The State of U.S. Science and Engineering 2022* (Arlington, VA: National Science Board, National Science Foundation, 2022), https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force.

8. 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.

9. Britta Glennon and David R. Dollar, “What’s Behind the Globalization of R&D?,” April 26, 2021, in *Dollar & Sense*, podcast produced by Brookings Institution, MP3 Audio, 25:02, https://www.brookings.edu/articles/whats-behind-the-globalization-of-rd.

10. Cecilia Esterline, “Previously Unreported Data: The U.S. Lost 45,000 College Grads to Canada’s High-Skill Visa from 2017 to 2021,” Niskanen Center, March 14, 2023, https://www.niskanencenter.org/previously-unreported-data-the-u-s-lost-45000-college-grads-to-canadas-high-skill-visa-from-2017-to-2021.

11. 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.

12. Clare McDonald, “Quarter of UK Businesses Looking Abroad for Tech Talent,” Computer Weekly, December 14, 2022, https://www.computerweekly.com/news/252528353/Quarter-of-UK-businesses-looking-abroad-for-tech-talent.

13. Michael Distefano et al., *The Trillion-Dollar Difference*, Korn Ferry Institute, 2016, https://www.kornferry.com/content/dam/kornferry/docs/article-migration/Korn-Ferry-Institute_The-trillion-dollar-difference.pdf.

14. 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.

15. 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.

16. 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* (Washington, DC: National Academies Press, 2014), https://doi.org/10.17226/18512.

17. James G. March, “Exploration and Exploitation in Organizational Learning,” *Organization Science* 2, no. 1 (1991): 71–87, http://www.jstor.org/stable/2634940.

18. Will Douglas Heaven, “Geoffrey Hinton Tells Us Why He’s Now Scared of the Tech He Helped Build,” *MIT Technology Review*, May 2, 2023, https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai.

19. American Chemical, “High Performance Carbon Fibers: National Historic Chemical Landmark,” September 17, 2003, https://www.acs.org/education/whatischemistry/landmarks/carbonfibers.html. The original publication is by Roger Bacon, “Growth, Structure, and Properties of Graphite Whiskers,” *Journal of Applied Physics* 31, no. 2 (February 1960): 283–90, https://doi.org/10.1063/1.1735559.

20. Congressional Research Service, *Federal Research and Development (R&D) Funding: FY2022*, January 2022, https://crsreports.congress.gov/product/pdf/R/R46869.

21. 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.

22. National Laboratories, “What We Do,” accessed September 19, 2023, https://nationallabs.org/our-labs/what-we-do.

23. National Research Council, *Research Universities and the Future of America: Ten Breakthrough Actions Vital to Our Nation’s Prosperity and Security* (Washington, DC: National Academies Press, 2012), https://doi.org/10.17226/13396.

24. National Academies of Sciences, Engineering, and Medicine, *Information Technology Innovation: Resurgence, Confluence, and Continuing Impact* (Washington, DC: National Academies Press, 2020), https://doi.org/10.17226/25961.

25. Elie Dolgin, “The Tangled History of mRNA Vaccines,” *Nature*, October 22, 2021, https://www.nature.com/articles/d41586-021-02483-w.

26. 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.

27. Mehdi Jedinia, “History of Assassinations of Iran’s Top Nuclear Scientists,” Voice of America, December 03, 2020, https://www.voanews.com/a/extremism-watch_history-assassinations-irans-top-nuclear-scientists/6199135.html.

12

TECHNOLOGY APPLICATIONS BY POLICY AREA

This chapter explores applications from each technology field described in the report as they may relate to five important policy themes: economic growth, national security, environmental and energy sustainability, health and medicine, and civil society. For each area, we extract from the technology discussions of chapters 1 through 10 applications or consequences that speak to it. Readers are invited to refer to the relevant technology chapter for more information about each application or consequence mentioned.

Economic Growth

Artificial intelligence 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 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 As much as 60 percent of the physical inputs to the global economy could be affected by biological processes. Biotechnology and synthetic biology are enablers for advances in medicine and health care, such as new vaccines and treatments for diseases including Alzheimer's, diabetes, and cancer. Synthetic biology also underlies advances in agriculture (e.g., drought-resistant crops), food (e.g., plant-based proteins), and energy production (e.g., biofuels). These advances could improve crop yields and boost energy production, lowering costs for consumers and solidifying US leadership in the field.

Cryptography Blockchain technologies can effectively provide provenance in supply chains and personal identity management that curbs fraud and identity theft, leading to more secure transactions and increases in seller efficiency. A US central bank digital currency could help reduce inefficiencies in US deposit markets, promoting broader participation in the financial system.

Materials science Lighter and stronger materials will increase the energy efficiency of vehicles used to transport people and cargo, leading to increased distribution of goods. New semiconductor materials are enablers for new types of chips and other information processing hardware. Technological advancements are also offering new ways to achieve low-carbon steel and cement production, which will help to reduce CO_2 emissions.

Neuroscience Neurodegenerative diseases—including chronic pain and subsequent opioid dependencies—currently lack effective treatments. Neuroscience is the best hope we have today for science-based interventions to reduce the symptoms and treat underlying conditions. While such interventions may be able to improve quality of life, they will also become increasingly important as the average age of citizens rises.

Nuclear technologies Nuclear-generated electricity is widely considered to be a necessary part of a net zero-emissions energy mix in the future. However, the lack of a US waste disposal policy is a substantial impediment to more widespread deployment of nuclear power in the United States.

Robotics Robots are used widely today, including in manufacturing, warehouse logistics, surgery, science and exploration, food production, disaster assistance, security and military services, and transportation. Advancements in robotics have enormous potential to affect jobs involving physical labor and presence.

Semiconductors Taiwan controls most of the world's production of semiconductors. To promote the US domestic semiconductor manufacturing industry, the White House signed into law the CHIPS Act in 2022. The act was also intended to incentivize companies to invest in American fabs. One year after the law was enacted, companies had announced $166 billion in investments, though many of those projects are contingent on the approval of federal aid.

Space Growth in the space sector is primarily driven by commercial and private activities and is expected to continue. Already, commercial space activities play critical roles in our daily lives and the economy. Satellites enable global navigation systems, guiding everything from autonomous cars to drones. Satellites also facilitate financial transactions, allow for more accurate weather predictions, and can even provide internet connectivity to people in remote, war-torn, or censored areas without broadband access.

Sustainable energy technologies Significant growth in the US renewable energy market is expected in the next several years as costs decline. In some local markets, renewable energy may become less expensive than fossil fuels. Widespread renewable energy deployment also has large macroeconomic effects. Doubling the share of renewables by 2030 would increase global GDP by over $1 trillion in addition to creating 24 million new jobs in the renewable energy sector.

National Security

Artificial intelligence Because AI enables more rapid processing of more data inputs, all aspects of military operations potentially benefit. Possible applications include managing military logistics, improving equipment maintenance effectiveness and efficiency, managing electronic medical records, navigating autonomous vehicles, operating drone swarms, recognizing targets, performing intelligence analysis, developing options for command decisions, and red teaming and war gaming to develop and refine plans.

Biotechnology and synthetic biology With synthetic biology becoming increasingly available to state and nonstate actors, many concerns arise that a malicious actor could create or deploy weaponized organisms or threaten the provision of biologically developed foods, medicines, fuels, or other products to coerce others.

Cryptography Adversaries are likely to have been storing encrypted data, and even though they were unable to read them at the time of storage, they hope future advances will allow them to crack the encryption. That future is the quantum future, and managing potential fallout from this scenario is a policy problem that will need to be faced when quantum computers come online.

Materials science Improvements in materials science and nanotechnology can improve capabilities in stealth, camouflage, and body armor, and can increase the energy content in explosives. Quantum dots, or materials that are smaller than about 100 nanometers in all directions, 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 posttraumatic stress disorder, thereby leading to better treatments for these conditions.

Nuclear technologies There are concerns that a global increase in fission reactors will result in a greater risk of nuclear proliferation, especially to current nonnuclear states or nonstate actors, while some believe that the emissions-free potential of fission reactors can minimize the risk of proliferation. The United States does not offer competitive exports of nuclear power plants; Russia, the United Arab Emirates, and South Korea lead this global market. The United States currently imports more than 90 percent of its uranium—about half from Kazakhstan and Russia and some 30 percent from Canada and Australia. Uranium extracted from seawater may decrease foreign dependence.

Robotics Advancements in robotics can assist US forces with load carrying, 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, also raise questions of roboethics 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 for semiconductor chips globally is Taiwan, which houses two of the three leading manufacturers (TMSC and UMC). China's long-held interest in reunification with Taiwan and its rising military capabilities and assertiveness toward Taiwan are raising deep concerns about the potential for a Chinese blockade or other actions that could disrupt the semiconductor supply chain for the United States and raise the risk of military conflict between the United States and China.

Space Communications, surveillance, and navigation in denied areas are essential functions for military forces. In the future, nonnuclear weapons may be based in space, for attack on terrestrial and/or space targets. Satellites are also essential for 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 is a driver of open-source (unclassified) intelligence, which has the potential to upend traditional intelligence processes built on classified information collection and analysis. The net effect of open-source intelligence 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.

Sustainable energy technologies The United States is no longer the world leader in energy manufacturing at scale; China and other countries with lower operating costs control most of the manufacturing, supply chain, and critical minerals for battery and solar cell production. Since these technologies will be directly tied to US energy security, it will be important to promote domestic production as well as collaboration with allies and partners to better protect energy supply chains.

Environmental and Energy Sustainability

Artificial intelligence AI capabilities can greatly improve global sustainability efforts, from helping farmers and hunters identify which produce or livestock are appropriate to harvest to helping analyze weather patterns to prepare populations and infrastructure for extreme or unusual conditions.

Biotechnology and synthetic biology Synthetic biology can contribute to new methods for energy production and environmental cleanup. It can also create more efficient fuel production, construction materials, and chemical processing; stabilize agriculture and aquaculture systems to address food scarcity; and improve food safety.

Cryptography Blockchain technologies can provide a transparent and secure way to track the movement of goods, their origin, quantity, and so forth, thereby improving efficiency in global supply chains and limiting underground or illegal extractions of certain materials.

Materials science Advancements in materials science and engineering are creating new and sustainable plastics that are easier to recycle. New materials design is also integral to decarbonization through electrification of transportation and industry. New materials will support the design of batteries capable of quick recharging, long stability, and cost reduction. Nanomaterials such as quantum dots can further improve the efficiency of solar cells and biodegradable plastics.

Neuroscience Sustainability on a planet with finite resources requires that decision makers and the people they represent be 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 fMRI brain scans of people making choices between immediate and delayed reward.[1] (This example is not further discussed in chapter 5.)

Nuclear technologies While nuclear power remains essential in the effort to decarbonize the energy industry, the capacity for nuclear reactors to generate electric power has declined in recent years, with new reactors coming online, mainly in Asia, unable to replace capacity loss from aging and decommissioned reactors in the West. It is unclear whether a sufficient number of nuclear reactors will become operational in time to reduce greenhouse gas emissions at a useful scale. Nuclear waste remains an environmental policy issue, and the United States has no enduring plan for a long-term solution to storing nuclear waste.

Robotics The deployment of robotics primarily for the Three Ds—dull, dirty, or dangerous jobs—enables robotic cleanup of environmentally hazardous materials and operation in environments inappropriate for humans, such as nuclear reactors.

Semiconductors Transitioning to renewable energy sources will require vast amounts of semiconductors. Advanced chips are integral to electric vehicles, solar arrays, and wind turbines. They are also used in smart devices and infrastructure that can self-monitor to consume energy more efficiently.

Space Satellite imagery can provide data on urban sprawl and global processes on land and at sea, including drought and ice cap melt, data that can inform sustainable development policies. In the next five years, there is room for development of space technologies 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 and also detect greenhouse gas emissions to identify natural-gas leaks and verify regulation compliance. Advancing space technologies can also enable mining of minerals from the moon or asteroids that are rare to find on Earth or transmission of sustainable solar energy directly to Earth from space.

Sustainable energy technologies The US government is investing in research-and-development projects across new energy technologies, enabling advancements in clean electricity generation, long-distance transmission lines, lighting based on light-emitting diodes, and improvements in electric car batteries. Long-duration energy storage is a critical field for climate and sustainability goals. Developing batteries for grid-scale storage across weeks or months are necessary to complement intermittent renewable energy generation. Hydrogen will power fuel-cell automotive vehicles and industrial processes. Currently, hydrogen is sourced from fossil fuels; sustainable hydrogen production methods that are cost-effective at scale are needed.

Health and Medicine

Artificial intelligence AI data analytics are already improving the accuracy of health-care assessments and procedures, and continued advancement in the field 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 new methods for pharmaceutical synthesis as well as pathogen detection and neutralization. Synthetic biology can additionally reduce disease transmission through gene drives, personalize medicine through genetic modifications, cure cancer with mRNA vaccines, and offer custom lab-grown human tissue for medical testing using “organoids.” DNA sequencers and synthesizers using the internet allow researchers to move viruses (and potentially vaccines or cures) around the world even faster than a pandemic. However, that same speed and accessibility creates concern for misuse by bad actors. It is also unclear how some new biological organisms will interact with the natural and human environments.

Cryptography 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. Such data storage can allow data analytics to be performed on aggregated and unassociated datasets, thus enabling researchers and internal auditors to access the needed information without violating patients’ privacy rights.

Materials science Materials science and nanotechnology are improving the abilities and effectiveness of medical devices and delivery. For example, wearable electronic devices made from flexible materials can conform to skin or tissues to provide specific sensing or actuating functions; devices like “e-skin” can sense external stimuli such as temperature or pressure; and “smart bandages” with integrated sensors and simulators can accelerate healing of chronic wounds by 25 percent. Injectable hydrogels can fine-tune long-term delivery of medications, which can lead to improvements in administration and the efficacy of essential medicines such as insulin. Nanomaterials like quantum dots are using fluorescent markers in biological systems to improve the contrast of biomedical images.

Neuroscience Advancements in neuroscience can address neurodegeneration and related diseases, such as chronic pain, opioid dependency, or Alzheimer’s, dramatically improving the quality of life and potentially reversing the anticipated rising costs associated with care. The annual cost of Alzheimer’s, for example, is projected to reach $1 trillion by 2050.

Nuclear technologies Medical isotopes, often radioactive, are used to diagnose and treat conditions such as heart disease and cancer.[2] Medical isotopes are often produced in nuclear reactors, although not reactors designed to generate electricity. (This example is not further discussed in chapter 6.)

Robotics Some robotics are already deployed in the health-care industry, such as assisted laparoscopic surgical units and equipment. Improvements in haptic technology (which gives the user a sense of feel—such as a smartwatch vibrating when a text is received) can increase the effectiveness and safety of these tools while providing new capabilities like soft and wearable robotic technologies. Robots can also help nursing and home-care workers provide essential functions such as bathing or cleaning.

Semiconductors Semiconductor chips are ubiquitous in modern medical equipment. Imaging devices such as MRI, CT, and ultrasound use embedded computers to generate images from electromagnetic radiation and sound waves penetrating or emanating from the human body. Tiny wearable health monitors and ingestible micro-robots would not be so small without embedded chips. Precision robotic surgery would not be possible without digital-to-analog converter chips. Across these examples, innovation in chip design, new materials, and integration methods are enablers for the performance, size, and efficiency of medical devices.

Space The potential for space manufacturing can improve development of specialized pharmaceuticals by utilizing the space environment—a very clean microgravity environment with minimal contaminants.

Sustainable 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 could prevent roughly fifty thousand deaths and save billions of dollars per year. Reduction of CO_2 emitted into the atmosphere will result in less extreme climates, which in turn will lead to fewer health problems from extreme heat.

Civil Society

Artificial intelligence Because AI models are trained on existing datasets, they are likely to encode any biases present in these datasets. This leads to inherent bias in AI and large language model systems, which can, in turn, affect decision making or model-based outcomes. For example, research has found that many facial recognition algorithms are better at identifying lighter-skinned faces than darker-skinned faces because of the training data used to develop them. This performance difference has led to cases of wrongful arrest of African Americans. AI models are also poor predictors of discontinuous change.

Biotechnology and synthetic biology Different religious traditions may have different stances toward life or living systems and whether the engineering of new life forms violates any of their basic precepts.

Cryptography The nature of cryptography and encrypted communications leads to some debate on exceptional access. Exceptional access regulations would require communications carriers and technology vendors to provide access to encrypted information under specific legal conditions, because the technology of encryption is accessible to criminals and other malefactors. Opponents of exceptional access argue that implementing this capability weakens the security provided by encryption. Supporters of exceptional access argue that lower personal encryption security is worth the benefits to law enforcement.

Materials science As with regulation in other areas of technology, concerns arise about the appropriate balance between promoting public safety from possible downside risks and the imperatives of innovation to move fast and leapfrog possible competitors. US ability to lead in this field is dependent on pathways for foreign talent to gain permanent residence, especially for PhD and advanced-degree graduates.

Neuroscience Cognitive and behavioral neuroscience also has broad implications for public policy, in that a basic aspect of criminal law is the nature and extent of an individual's responsibility for a criminal act. Minors under eighteen years of age, for example, cannot be subject to the death penalty for crimes they committed, because adolescent brains are not fully developed, putting minors at higher risk of impulsive, irrational thoughts and behaviors.

Nuclear technologies The construction of nuclear power plants or facilities for storing radioactive waste is often met with opposition from those concerned about exposure to radiation in the environment.

Robotics A challenge in designing human-robot interaction (HRI) is unpredictable or unintended physical contact that can cause safety issues for the human. Another challenge is to design HRI in a way that accommodates social norms and that allows robots to exhibit behaviors that are more familiar and comfortable for humans. As robots assume more roles with decision-making components, some concepts of individual accountability may be challenged and need to evolve.

Semiconductors Student interest in hardware design has dropped precipitously in favor of software-oriented jobs. Some estimates suggest that the semiconductor manufacturing employment sector will only be able to fill 30 percent of its needs by 2030.

Space As public entities grow more risk-averse and private companies more risk-tolerant for the sake of financial gain and innovation, collaborative efforts between academia and industry are pivotal for continued leadership and development in this field. In space, rapid expansion driven by increasing commercial assets and applications is exceeding the existing policy context for space activities.

Sustainable energy technologies Continued creation of sustainable energy infrastructure requires new acquisitions of land to build generating stations and storage facilities, which can displace residents from private property and impact local property values (i.e., we support windmills but "not in my backyard"). Additionally, US energy policy is governed by a variety of overlapping federal, state, and local government jurisdictions that can complicate new energy initiatives or incentives.

NOTES

1. Emmanuel Guizar Rosales, Thomas Baumgartner, and Daria Knoch, "Interindividual Differences in Intergenerational Sustainable Behavior Are Associated with Cortical Thickness of the Dorsomedial and Dorsolateral Prefrontal Cortex," *NeuroImage* 264, no. 119664 (2022), https://doi.org/10.1016/j.neuroimage.2022.119664.

2. Institute of Medicine, *Isotopes for Medicine and the Life Sciences* (Washington, DC: The National Academies Press, 1995), https://doi.org/10.17226/4818.

CONCLUSION

The ten technology areas highlighted in this inaugural issue of the *Stanford Emerging Technology Review* represent some of the most important fields in science and technology on the horizon today. After interviews with seventy-five leading faculty members across thirty scientific disciplines, it is clear that the following decade will see the convergence of multiple technologies to drive progress at unprecedented rates. Artificial intelligence, driven by increasing computing power and data, will boost human productivity to unprecedented levels and supercharge progress in almost every field. Synthetic biology and biotechnology offer revolutionary capabilities for agriculture, medicine, and even manufacturing. Progress is happening in the vast region of space all the way down to nanoparticles smaller than a human hair. While technology is inherently neither good nor bad, it is important for policymakers to grasp the magnitude of technological change, the potential for these emerging tools to improve or exacerbate societal norms—and the necessity of American leadership in navigating the ever-expanding frontiers.

For many decades, the mind-set of US science and technology policy has been to fund research at universities and national labs, wait for breakthroughs, and hope they bring about positive change. This moment requires a different approach. Policymakers must engage more strongly with the technology community, in both academia and the private sector, to shape the ecosystem in a way that serves and protects the interests of the American people.

Ultimately, technology is developed and used by humans, and effective governance to multiply benefits and reduce risks requires human direction. Policymakers can develop frameworks that foster innovation, set priorities and strategies, align economic strategies for innovation and continued leadership, and strengthen the hand of the United States where it needs to compete internationally. Developments in renewable energy, for example, offer a path to energy independence and sustainability, but they will require sustained government attention and funding in order to attract innovators and overcome the valley of death.

This report offers a state-of-the-union look at ten technology fields: artificial intelligence, biotechnology and synthetic biology, cryptography, materials science, nuclear technology, neuroscience, robotics, semiconductors, space, and sustainable energy technologies. We outline the most important developments and advances, explain key technical details to the layperson, and offer a view of the critical policy considerations governments will have to debate over the coming years. Based on our research, we highlight common themes that emerge across the ten technology fields 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 key point. Synergies between different technologies are large and growing. Worldwide, access to technologies is growing as they ride a declining cost curve, supplemented with innovations in portable laboratories and open-access resources.

Gaining a lead in a technology area is not the same as sustaining it. Engaging with expertise around the world, tapping the potential of highly skilled and educated immigrants, and maintaining robust domestic pathways to scientific expertise will bolster American leadership in an increasingly competitive global environment. Recognizing the changing role of government in technology development is also critical. No longer are innovations solely developed and guarded by state-backed research groups; private corporations and even individual actors are able to develop transformative technologies.

This paradigm shift is most evident in fields like artificial intelligence and space, where private companies are leading the charge on large language model systems and placing smaller and more advanced assets into space, a field formerly reserved for governments. Concentration of power in different hands has important implications for technology access, priorities, and policy.

We began this report by asking the question "What do policymakers need to know about emerging technologies from Stanford?" This report is the first step in answering that question, by providing the knowledge necessary for these crucial technologies, their key takeaways, future implications, and potential policy concerns to foster meaningful and enduring conversations that can lead to appropriate, supportive, and timely legislation even as technologies continue to change.

In the months ahead, SETR will be producing deeper-dive reports on the ten technological areas, holding briefings in California and Washington, DC, and launching multimedia educational products.

Gaining a lead in a technology area is not the same as sustaining it.

CONTRIBUTORS

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 the nineteenth national security advisor (2001–5). She is author and coauthor of nine books. Her research interests include education, security and defense, US foreign policy, law and policy, politics and public opinion, and the US presidency. She received her PhD in political science from the University of Denver.

John B. Taylor is the George P. Shultz Senior Fellow in Economics at the Hoover Institution and the Mary and Robert Raymond Professor of Economics at Stanford University. He is also the director of Stanford's Introductory Economics Center. He specializes in monetary policy, fiscal policy, and international economics. He served as undersecretary for the Treasury for international affairs (2001–5) and on the President's Council of Economic Advisers. He has authored and coauthored twenty books, including *First Principles: Five Keys to Restoring America's Prosperity*, winner of the 2012 Hayek Prize. He received his PhD in economics from Stanford University.

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 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 of political science (by courtesy) 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.

Director and Editor in Chief

Herbert S. Lin is senior research scholar at the Center for International Security and Cooperation and Hank J. Holland Fellow in Cyber Policy and Security 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.

Faculty Council

Zhenan Bao is the K. K. Lee Professor in Chemical Engineering and professor, by courtesy, of chemistry and materials science and engineering at Stanford University. She has close to seven hundred refereed publications and more than eighty US patents. Her current research focuses on organic electronics, including skin-inspired materials, dynamic energy storage, and recyclable, reprocessable materials. She received her PhD in chemistry from the University of Chicago.

Dan Boneh is 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 one hundred publications. He received his PhD in computer science from Princeton University.

Yi Cui is the director of the Precourt Institute for Energy, the Fortinet Founders Professor of materials science and engineering, and professor of energy science and engineering and, by courtesy, of chemistry at Stanford University. He is also a senior fellow at the Stanford Woods Institute for the Environment and codirector of Stanford StorageX Initiative. His research focuses on nanomaterials, nanobiotechnology, and nanoenvironmental technologies. He received his PhD in chemistry from Harvard University.

Simone D'Amico is associate professor of aeronautics and astronautics and professor, by courtesy, of geophysics at Stanford University, where he serves as the W. M. Keck Faculty Scholar of Engineering. His research explores the intersection of advanced astrodynamics, spacecraft navigation and control, autonomous decision making, and space system engineering. He currently leads 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 Fellow in Undergraduate Education (bioengineering), codirector of degree programs for the Hasso Plattner Institute of Design (the d.school), core faculty at the Center for International Security and Cooperation (CISAC), and senior fellow (courtesy) of the Hoover Institution at Stanford University. He serves as president and director of the Biobricks Foundation and 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 earned a PhD in biotechnology and biochemical engineering from Dartmouth College.

Siegfried Glenzer is professor of photon science and, by courtesy, of mechanical engineering at Stanford University, where he serves as director of the High Energy Density Science division at SLAC National Accelerator Laboratory. His research explores matter in extreme conditions and the development of fusion energy sciences and technologies. He has authored or coauthored more than 550 journal publications. He received his PhD in physics from Ruhr University Bochum.

Mark A. Horowitz is the Yahoo! Founders Professor in the School of Engineering and professor of electrical engineering and computer science 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.

Fei-Fei Li is the Sequoia Professor of Computer Science and professor, by courtesy, of psychology at Stanford University. She serves as codirector of Stanford's Institute for Human-Centered AI and as an affiliated faculty at Stanford Bio-X. Her current research includes cognitively inspired AI, machine learning, computer vision, and ambient intelligent systems for health-care delivery. She received her PhD in electrical engineering from the California Institute of Technology.

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 a deputy director of the Wu Tsai Neurosciences Institute, affiliated faculty at Stanford Bio-X and Stanford's Institute for Human-Centered AI, 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 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.

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. Gordon is currently the director at CACI International and served as the fifth principal deputy director of national intelligence at the Office of the Director of National Intelligence (2017–19). She was previously deputy director of the National Geospatial Intelligence Agency (2015–17) and served twenty-seven years in the CIA, including as director of the CIA's Information Operations Center.

John L. Hennessy is 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 graduate-level scholarship program. He received his PhD in computer science from Stony Brook University.

Jerry McNerny 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 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, Nextdoor, 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, 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. and a member of the firm's operating committee and cofounder and chairman of the board of directors of the Greater Washington Partnership. Prior to joining JPMorgan Chase in 2008, he was the managing partner of the Washington, DC, office of law firm Mayer Brown LLP and spent nearly a decade in public service.

Eric Schmidt is cofounder of Schmidt Futures, a philanthropic initiative that bets early on exceptional people making the world better, and chair of the board of trustees of the Broad Institute. Previously, he was CEO and chairman of Google (2001–11) and executive chairman and technical adviser to Alphabet until 2020. 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 at Urbana-Champaign and the University of California–Berkeley and is an elected member of the American Academy of Arts and Sciences.

Stanford Emerging Technology Review Fellows

Alaa Eldin Abdelaal, robotics
Jesus Miguens Blanco, biotechnology and synthetic biology
Stefano Cestellos-Blanco, materials science
Alec Friedrich Condon, neuroscience
Sneha Goenka, semiconductors
Tommaso Guffanti, space
Melanie Hannebelle, biotechnology and synthetic biology
Nicholas Hartley, nuclear technologies
Sang Cheol Kim, energy
John Klich, biotechnology and synthetic biology
Nicolas Lee, space
Hope Tianfeng Leng, biotechnology and synthetic biology
Daniil Lukin, semiconductors
Walter Joshua Manuel, space
Lukas Felix Michalek, materials science
Neil Perry, cryptography
Tathagata Srimani, semiconductors
Artem Trotsyuk, biotechnology and synthetic biology

Senior Research Staff

Ashil Jhaveri
Taylor McLamb
Alexandra Sander
Corie Wieland

Student Research Assistants

Mandy Alevra
Lily Barnett
Mishika Govil
Jasdeep Singh Hundal
Nicole Jones
Dwight Knightly
Jennifer Lin
Reid McCaw
Antonio Noyola
Isobel Douglas Porteous
Arman Sharma

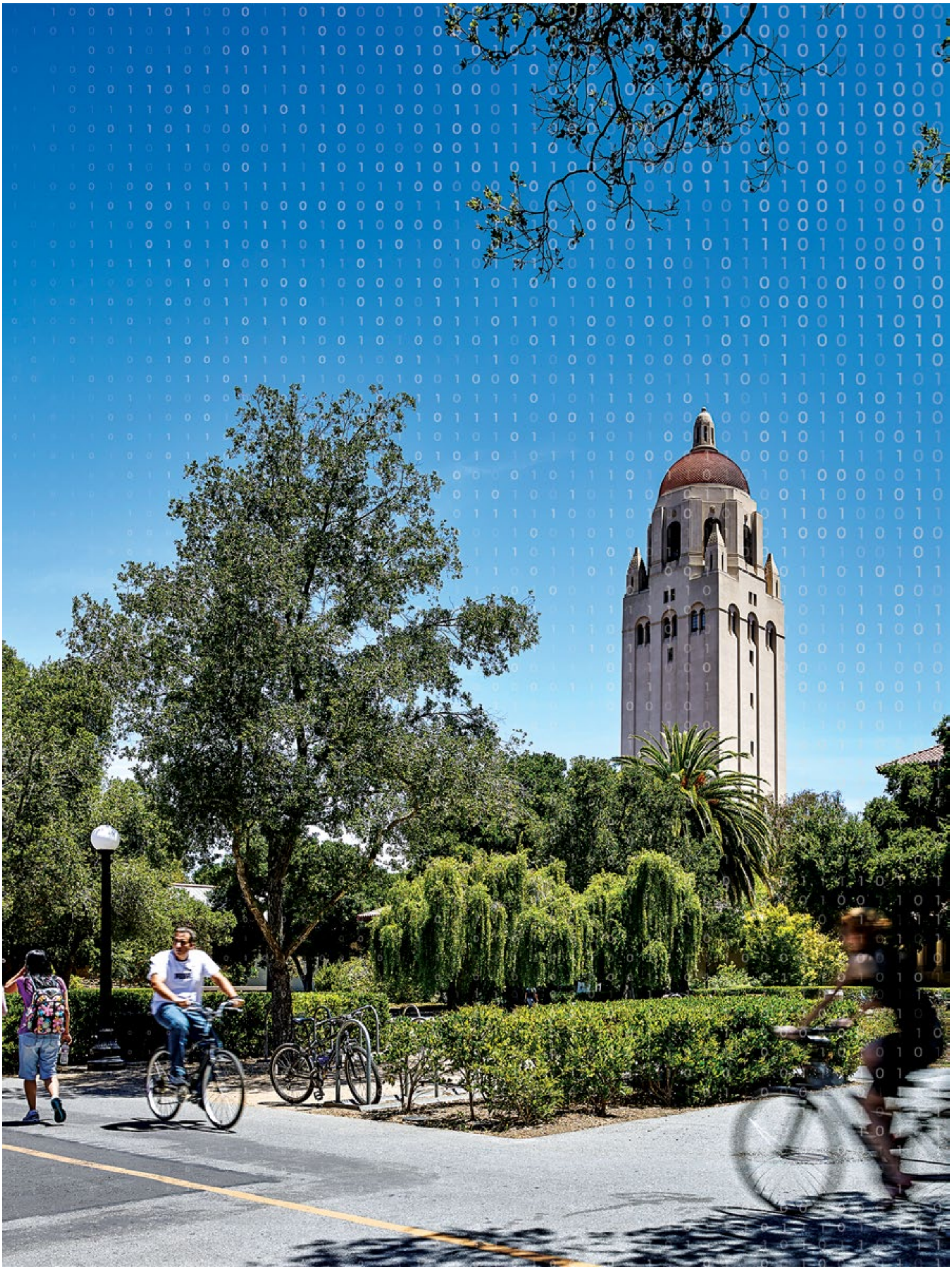
---

29 28 27 26 25 24 23   7 6 5 4 3 2 1

Cover and facing photos by Tim Griffith

Designer: Howie Severson

Image credits: iStock.com/mofuku (p. 20); iStock.com/wacomka (p. 32); iStock.com/FeelPic (p. 44); iStock.com/Chartchai Sansaneeyashewin (p. 54); iStock.com/ArtemisDiana (p. 64); iStock.com/peterschreiber.media (p. 74); iStock.com/imaginima (p. 84); iStock.com/Floriana (p. 96); iStock.com/dima_zel (p. 106); iStock.com/JONGHO SHIN (p. 118)



NOTES