1 From Inventions to Systems

This book is about what the world humans have created has become. It is neither a history of technology or inventions, although both factor into our story, nor a social history, although the social aspects of our world are inextricably linked to our story. Ultimately, it is about the challenges our world presents to new generations and how these are being met—and will be met in the future—through new ways of thinking. Though our focus is on what has traditionally been called "engineering," the challenges are for us all.

Engineers are justly proud of the many amazing inventions for which they are responsible, and for the classical methods they and their "forefathers" have created for designing, building, and operating the systems of the mid- to late twentieth century. But are these older approaches that were fine for, say, Henry Ford's Model T, sufficient for the automobile of the future? What about the intertwined social and technological complexities of today, and of the years to come? Can we couple the social dynamics, economics, and technological evolution of human-made systems and figure out what works best, what needs to be done, and how to make it all happen?

The answers lie in a world where the clear boundaries between the old systems are blurred, and the blurring is not feared but embraced. It is the world of *engineering systems*, where scholarship and knowledge at the nexus of engineering, management, and the social sciences is making it possible to identify the challenges and meet them, identify the opportunities and capture them, and prepare for the unknown that awaits us beyond.

To gain an understanding of those challenges and how they are approached differently in an engineering systems context requires looking back at some of the milestones in human history that created the world around us today.

The Genius of Invention

We humans have always sought to shape our environment to suit our needs. Beginning with early tools, we have created objects specifically designed to make it easier to perform the basic tasks of survival, such as hunting and cultivating food and constructing shelter, and to enhance the quality of our lives, for example, by creating stimulating social interactions. This has been true throughout our long history, but with the Industrial Revolution of the eighteenth century and the "great inventions" of the nineteenth century we entered an era of continuous explosive growth and innovation.

Since then, the inventions have been coming at an accelerating rate. The invention of the automobile, often attributed to Karl Benz in 1885, gave us unprecedented personal mobility. The telegraph (attributed to Samuel F. B. Morse in 1837) and telephone (Alexander Graham Bell in 1876) enabled the fast transmission of messages over long distances. The incandescent light bulb of Thomas Edison (1879) helped turn night into day. Each of these advances emerged from a complex tapestry of experimentation, failure, more advances, and ultimate success woven over long periods of time.

In 1600, nearly 250 years before Thomas Edison was born, an English physician, William Gilbert, distinguished the lodestone effect produced from static electricity by rubbing amber, and—referring to its property of attracting small objects—coined the new Latin word *electricus* ("of amber") from the Greek *elektron* ("amber"). It was another half-century before the words "electric" and "electricity" appeared in print, and nearly an additional full century before, in 1733, the French chemist Charles François Cisternay du Fay published his discovery of two types of electricity, positive and negative charged. And all this took place before Benjamin Franklin flew his famous kite in 1752. By 1800, Alessandro Volta had made the first electrolytic cell—the basis for the battery.

Meanwhile, humans were expanding their transportation options. In 1801, Richard Threvithick built a steam-powered road locomotive, and 5 years later François Isaac de Rivaz invented the first internal combustion engine, fueled by a mixture of hydrogen and oxygen. In Germany, a physician and avid inventor named Samuel Thomas von Sömmering was thinking not about personal mobility, but about how to send information from one place to another without using the postal service. In 1809, he created an electrochemical telegraph that could convey messages electrically over distances of up to a few kilometers. His achievement came to be seen as an advance not in the field of electricity, but in communications, just as Michael Faraday's 1821 invention of the electric motor was a milestone in transportation.

Benz, Morse, Bell, Edison: these may be the names schoolchildren learn, but many individuals and institutions contributed, either directly or indirectly and cooperatively or uncooperatively, to each of their inventions, sometimes with only limited cross-fertilization. Many of these inventions were even created in parallel in different parts of the world at roughly the same time.

At first, these innovations were seen as curiosities, as exotic artifacts. They were familiar primarily to those "in the know" about the latest advances in physics and the materials sciences, and were affordable only to a small, wealthy elite. Their full value was only realized when they were made robust and connected through larger networks and infrastructures and became commmonly used by the masses. Further technical changes, such as those that expanded the food supply and enabled a longer lifespan, accelerated world population growth—which in turn raised new challenges.

Inventions Begin to Be Connected

In 1750, the Earth had an estimated 791 million inhabitants; North America had a mere 2 million. A hundred years later, those numbers had grown to 1.26 billion and 26 million, respectively. A decade or so later, accelerated population growth was assured by the Frenchman Louis Pasteur, one of the fathers of microbiology, whose experiments confirmed the correctness of the germ theory of disease. His research showed the role that microorganisms play in spoiling milk, beer, wine, and other beverages, and his invention of the pasteurization process to kill the bacteria that caused so much sickness in humans tremendously boosted human longevity. Pasteur's contributions to immunology and the creation of vaccines further improved health and accelerated the growth in world population, as shown in figure 1.1.[1] Of course, much of this growth resulted from breaking the "Malthusian Trap" through the increased efficiency of the food production system enabled during the era of great inventions.[2]

Around the same time, Antonio Meucci was constructing the first electromagnetic telephone to connect his second-floor bedroom to his basement laboratory in Staten Island, New York. There were certainly a lot more people around to benefit from such an invention. By 1950,

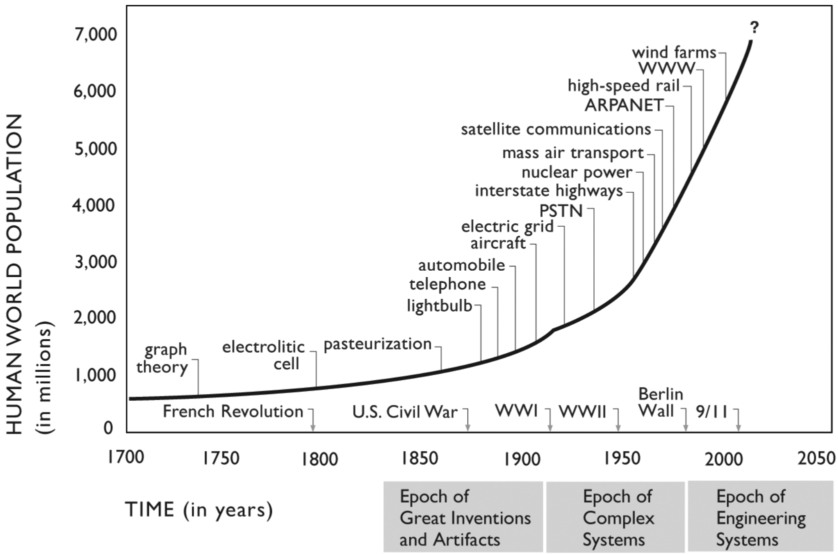

Figure 1.1
Human population growth and important technological and political milestones (1700–2050).

the world's population had grown to 2.52 billion, and inventions were "growing" along with the population. Consider only the telephone and telegraph: New Haven, Connecticut, got the first telephone exchange in 1878, and in the 1880s, the public switched telephone network (PSTN) emerged in the United States. The first long distance connection, from Boston to New York City, was created in 1884. By 1902, the telegraph cables encircled the entire world, and a decade later long-distance ranges for the telephone had reached Denver, the technological limit. Three years after that, Alexander Graham Bell in New York City was able to phone Thomas Watson in San Francisco. The first wireless message was transmitted across the Atlantic Ocean in 1901 by Marchese Guglielmo Marconi.[3]

From then on, communication capabilities were being improved with the new technology that followed from Marconi's work—radio. By the end of 1928, a fascinating 30-year-old Slovenian born in what is now Croatia, Herman Potočnik, had published a groundbreaking book in which he calculated the geostationary orbit of an imagined space station and described communication between the station and the ground using radio.[4] The book featured 100 quite amazing handmade illustrations, and was the precursor of Arthur C. Clarke's notion, detailed in 1945, of three geostationary satellites for mass communications that would provide global coverage. Meanwhile, back on Earth, the mayor of Englewood, New Jersey, phoned the mayor of Alameda, California, making the first long-distance telephone call that did not require the assistance of an operator.

All these people who were communicating were also beginning to drive cars. Population growth and increased technical capability made the car desirable to more and more people, and in 1902 Ransom Olds, who had been tinkering with automobiles and their engines for years, debuted large-scale, production-line manufacturing of affordable cars. The earlier evolution of the train network had been a major engine of westward expansion in the United States.[5] Henry Ford stood on the shoulders of Olds, in 1913, when he created the Ford assembly line. As affordable cars became accessible to the growing populations of the United States and Europe, governments began to think about the transportation infrastructure. During the Weimar Republic of the 1920s, the Germans conceived of a national highway system, and in 1921 the U.S. Army was asked to provide a list of roads it considered necessary for national defense—the precursor to a nationwide highway system in the United States. New England had already established its own network of "interstate" roads in 1922.[6]

Meanwhile, automobile manufacturers had begun to think beyond the technological aspects of the car as an invention and consider the business side of the equation to a far greater degree. Merging his roller and ball bearings company with the company that eventually became General Motors, Alfred P. Sloan rose through the firm's executive ranks. As GM's president beginning in the 1920s, Sloan introduced product differentiation and market segmentation, with a pricing structure that avoided competition within the GM family of cars and kept consumers buying from the company even as their income grew and preferences evolved. He established annual styling changes—an idea that led to the concept of planned obsolescence. He adopted from DuPont the measure of return on investment as a staple of industrial finance. Under Sloan, GM eclipsed Ford to become the world's leading car company, as well as the world's largest and most profitable industrial enterprise for a long period.

Years later, GM's leadership—indeed, that of the entire U.S. automobile industry—would be challenged by Toyota and its Toyota Production System (TPS), an idea hatched by an engineer named Taiichi Ohno and supported by Sakichi Toyoda and his son Kiichiro Toyoda. TPS organizes manufacturing and logistics, includes interactions with suppliers and customers, and represents a fundamentally different logic and framework than mass production for the business of developing, making, and selling cars. Most important, TPS was conceived of as an evolving system, not a "breakthrough" invention.

These Toyota founders visited America in the 1950s to see how the Ford assembly line worked, but left unimpressed by the large inventories, uneven quality of work, and large amount of rework required before a Ford car was truly "done." They found their inspiration, instead, at a Piggly Wiggly supermarket, where they saw how goods were reordered and restocked only after the store's customers had bought them. The rest is history—and notable not only because Toyota shook the auto manufacturing world with its approach, but also because the company directly challenged American and European car makers as the emerging global economy made it easier for Toyota first to sell its "better-made" cars worldwide and, eventually, to build them globally as well. Every global auto company was forced to rethink not only the underlying technology of the car, but also how it managed automobile research and development and car-building processes.

Networks and Infrastructures

The permeation of autos, telephones, and the electric grid that provided power to street lamps and homes fed the growth of networks and infrastructure. However, these networks (roads, power, and communications) were either not connected or only loosely coupled to each other, each doing its job more or less in isolation. Railroads and telegraphs provide a compelling example. Telegraph wires largely ran parallel to the railroad tracks that were making their way across North America, and the railroads used telegraph services to communicate between stations and depots along the rail lines, but these were separate networks that happened to be proximal to each other because it minimized costs.

Still, there were demands for more "services" to accommodate the invented artifacts available to a growing populace. Ford's manufacturing and distribution systems had put cars in the hands of American drivers, and it wasn't long before they were demanding more and better roads. The first "national road system" was completed in the 1920s and 1930s and, by 1956, construction was underway in the United States to create an interlocking network of interstate highways, exclusively for travel (absent commercial enterprises along the roads), with limited on-off access. This would have a transformative effect on the country. As telephones were connected first locally, then nationally, and finally internationally, networks were needed to ensure that demand for communication capabilities could be met. The electric light bulb and an increasing number of industrial and residential machines needed to be powered continuously, and the answer was to create national and transnational electrical energy grids for electricity production, transmission, and distribution.

In parallel with the growth of physical infrastructures came an increase in the size of the firms and organizations that ran these systems. The Pennsylvania Railroad Company, Standard Oil, "Ma Bell," and others became behemoths, and despite the economies of scale gained by their size serious concerns soon emerged about concentration of power and monopolism—and, in some cases, society stepped in to force changes.

What had happened was a metamorphosis. Demographic changes throughout the Western Hemisphere, in Europe, and in Asia were fueling the growth of something bigger and more complex than the world had ever seen. To be sure, the artifacts that had been invented decades earlier—the telephone, the automobile, the light bulb, the radio, the telegraph—were themselves complex, but they were self-contained. Now, they had become part of something else. As Thomas Edison understood, the light bulb was part of a technical energy *system* that included the power plants generating the electricity, the transmission lines bringing it to the bulb, and the coal mines supplying the fuel for generation. And, by the way, that coal got to the power plants on trains that were part of the larger transportation system, and which also were delivering automobiles from factories to those coal miners who, over time, were able to afford them. At the same time, the beginning of a multimodal communications system became apparent with the emergence of radio and television, which began to complement the telephone and telegraph systems and the physical delivery of newspapers, letters, and parcels.

Those descriptions barely begin to touch on the complexity of these systems in our day and age. Over time, the infrastructures required standards and regulations; some were established cooperatively, while others required government involvement. The industries within the systems needed materials from other industries, making, for instance, a direct link between the mining of iron ore on Minnesota's Mesabi Range, the production of steel in Pittsburgh and Chicago, and the creation of automobile bodies in Detroit factories. All these transactions were enabled by the improving communication system.

Unintended Consequences

Although some aspects of these growing systems were designed, to the degree they *were* designed, for single purposes and with a clear mission in mind—for instance, transporting people or transmitting messages—a lot of their present-day effects were essentially left to chance. Take, for instance, the traffic jam, something to which none of the early developers gave any apparent thought.

On July 11, 1910, the headline in Jacksonville, Florida's daily newspaper, the *Florida Times-Union and Citizen*, announced something the small city had never seen: "Autoists Spending Day At The Beach All Made Rush For The City At The Same Time!" The subhead described how, at the ferry crossing that linked the city with the new paved highway (the first in the southeast United States) that went to the beach: "Upwards Of 50 Cars Were Waiting At One Period!" A year later, on June 25, 1911, the same newspaper reported: "The constantly increasing number of automobiles in use in Jacksonville makes their safe navigation of the streets a more difficult problem in proportion. Hundreds of motorcars are using the streets every hour of the day and far into the night. In most cases they are left to work out their own salvation."[7]

Traffic jams were assuredly not the only unintended consequence of a great invention. Thomas Edison gave nary a thought to whether producing the electricity for his incandescent bulbs would result in air pollution. And beyond that, little or no consideration was given in the early days of these systems as to whether they would be stable and sustainable over the long term. In fact, the general mindset in the decades immediately following World War II was that resources were, for all intents and purposes, essentially inexhaustible. Smoke could be seen spewing from the stacks of factories, but these emissions were regarded as negligible and even as a sign of real progress—as evidenced by the artwork and photographs in many corporate headquarters proudly depicting factories billowing huge amounts of smoke.

Things changed when many systems reached a critical size or "tipping point." While component technologies continued to evolve rapidly—faster computers, better cars, safer aircraft, and so on—the underlying infrastructure networks that had formed, and especially the regulatory frameworks, stagnated, failed to anticipate changes, or simply did not keep up with growth. This mismatch between technological progress and the backwardness of infrastructures and regulations persists to some degree today.

Eventually, unintended consequences could no longer be ignored. Many of the most dramatic changes began in the 1960s—no doubt fueled in part by a younger generation coming of age after the "complacency" of the 1950s that viewed the world quite differently from their parents. We'll leave it to the political scientists to explain that, but we can point to several specific events.

In 1962, a marine biologist by the name of Rachel Carson, who in the 1950s had published a trilogy of bestsellers touching on nearly every aspect of ocean life, stunned the American public with a new book titled *Silent Spring*.[8] Carson had long been a conservationist with a particular interest in environmental hazards of synthetic pesticides. A U.S. Department of Agriculture program in 1957 to eradicate fire ants had turned Carson's interest into a crusade. She studied the aerial spraying of the pesticide dichlorodiphenyltrichloroethane, better known as DDT, and eventually worked with the Audubon Society to oppose such spraying. *Silent Spring* showed, with gripping examples, the environmental damage that DDT and other pesticides—which Carson dubbed "biocides"—were causing.

Silent Spring caused quite an uproar. Serialized in *The New Yorker*, it provoked the wrath of the chemical industry and stirred the public. When the book was announced as the October monthly selection of the Book-of-the-Month Club, Carson promised it would be carried to "farms and hamlets all over [the] country that don't know what a bookstore looks like." Meanwhile, industry lobbyists and lawyers were working overtime trying to stop the book from coming out.

Rachel Carson's *Silent Spring* led, eventually, to a nationwide ban on DDT and other pesticides. Her work is widely credited with spurring the creation of an environmental movement in the United States, which in turn led to the establishment of the U.S. government's Environmental Protection Agency. As her biographer wrote, Rachel Carson "quite self-consciously decided to write a book calling into question the paradigm of scientific progress that defined postwar American culture"[9]; its central theme was the often negative impacts humans have on the natural world that they seek to harness with technology. It is not an overstatement to say that the world would never be the same.

Three years after *Silent Spring*'s publication, another book, *Unsafe at Any Speed*,[10] challenged how people thought about technology and the world around them. Written by Ralph Nader, a lawyer who had been working for then U.S. Assistant Secretary of Labor Daniel Patrick Moynihan, it was to become a model for consumer advocacy.

In his book, Nader questioned both the overall unwillingness of U.S. car manufacturers to invest in safety improvements for their vehicles and specifically their resistance to seat belts, which had been introduced only sporadically and met with little customer enthusiasm.

The book definitely got the automobile industry's attention. General Motors added considerably to the book's impact when it tried to destroy Nader's public image by spreading rumors about every aspect of his life, political and religious views, sexual behavior, and personal habits. GM even spied on the young author, and in 1966 GM's president, James Roche, was forced to appear before a subcommittee of the U.S. Senate and apologize publicly to Nader for his company's activity. Nader later used the money won in a successful lawsuit against GM for invasion of privacy to become a full-time lobbyist for consumer rights, helping promote the creation of the U.S. Environmental Protection Agency and passage of the Clean Air Act.

Nader's book opened a decades-long battle, ongoing still, between consumer advocates, the government (more or less on the side of consumers or industry, depending on the specific circumstances), and automobile manufacturers over car safety, design, and fuel efficiency as well as many other aspects of the car as an artifact and its many effects in the world.

Another important event that received significant attention was the formation of the Club of Rome, which attempted to address the resource scarcity category of unintended consequences. In April 1968, Italian industrialist Aurelio Peccei and Scottish scientist Alexander King brought together a small group of academics, diplomats, industry leaders, and others at a meeting in Rome to discuss "the predicament of mankind." The release of the Club of Rome's report, *The Limits to Growth*,[11] in 1972 most definitely touched a nerve. The report, which has sold more than 30 million copies, promulgates the idea that economic growth cannot continue indefinitely because natural resources—particularly oil—are limited.[12]

The degree to which natural resources are, indeed, limited and whether more recycling, new materials, renewable sources of energy, and other technological innovations and regulatory responses will be able to shift or entirely overcome any real or perceived limits to growth are the subject of intense debate to this day. It is clear from today's perspective that the oil "crisis" of 1973 and the subsequent political and economic events of the 1970s caused a fundamental shift in the public's awareness regarding wide-ranging issues such as the availability of natural resources, the management of emissions and waste products, and the general belief that technological progress by itself can be the answer to society's problems. Awareness of these wide-ranging issues is, in fact, awareness of unintended consequences.

The story of asbestos illustrates another category of unintended consequences—*the potentially harmful side effects of technology*. It is a story that goes all the way back to the ancient Greeks, who called this naturally occurring silicate mineral the "miracle mineral" because it could withstand heat and was so easy to use. Its name comes from the ancient Greek word for "inextinguishable." Charlemagne is even said to have had a tablecloth made from it—perhaps the ultimate trivet.

When the Industrial Revolution came along, asbestos use widened. By the 1860s, U.S. and Canadian builders were insulating homes and other buildings with the material, either directly or by mixing its fibers with cement. It was wrapped around the wiring of electric ovens and hotplates. Eventually, it showed up in automobile brake pads, shoes, and clutch discs, protecting them from "burning out." Asbestos was even a component of the first filtered cigarette in 1952. The list goes on and on.

The problem is the unintended consequence of the technologies that made asbestos use easier and more economically viable: it is *deadly*. Early in the last century, researchers began to notice a growing number of lung problems in towns where asbestos was mined, and even a lot of early deaths. The first documented death from asbestos was in 1906, and diagnoses of *asbestosis* first appeared in the 1920s. No one knows what percentage of the human race may have died from the various diseases asbestos causes since the time of the ancient Greeks. To illustrate the *potential* number, consider one estimate that upward of 100,000 people may have died from exposure to asbestos in just one industry—U.S. shipbuilding. World War II-era ships had asbestos-wrapped pipes, asbestos-lined boilers, and asbestos-covered engines. Of the roughly 4.3 million wartime shipyard workers in the United States, an estimated 14 of every 1,000 died of mesothelioma caused by asbestos![13] Not until the late 1980s and early 1990s did most of the world's developed countries begin to attack the problem in earnest.

It is not clear whether the asbestos story would have been different if engineers had studied biology in earlier years. The biological sciences, physiology, and medicine have been strongly connected to engineering only since late in the twentieth century. However, it is apparent that the impact on human health and well-being must take center stage when developing new artifacts, improving technologies, and designing large-scale systems. Concerns of this sort, when given proper consideration, have often led to better and more sustainable designs. Sometimes, though, social concerns are misused to delay or cancel beneficial projects. Whereas traditional technical analysis typically produces an answer acceptable to all parties, the inclusion of sociotechnical factors frequently leads to different stakeholders making widely different claims, which in turn can have a significant impact on the development of solutions and implementation.

Growing Systems Interactions

The revolution in social attitudes and dynamics continued to transform attitudes toward and perceptions of resource scarcity after the 1970s. Heightened awareness, fueled by an explosion in the information available to people on nearly any topic you can imagine, created greater pushback than ever about the harmful side effects of technology. Meanwhile, though, technology continued to progress, and systems became even more complex and capable of making modern life simultaneously easier and more challenging. In the final decades of the twentieth century, new alternatives began to emerge for the major functions that make ours a modern society. Personal mobility could be achieved more safely, and at less expense, than ever before—not only by car, but also by rail, air, and boat. Communications were no longer limited to the mail, telephone, and telegraph, but there were fax machines, satellites, e-mail, and, more recently, numerous services such as Skype and VOIP (Voice Over Internet Protocol) that rely on the World Wide Web and the Internet. Power plants that once used only coal to produce electricity could be powered by oil, natural gas, nuclear energy, water, wind, and solar radiation. Systems that had once been clearly separate began to interact more than anyone could have imagined. Some of these interactions were deliberately exploited while others were more opportunistic or even accidental. For instance, cars had long had radios, but now you could make a telephone call from your car. The role of humans in these systems as designers, funders, operators, and users became increasingly complex and multifaceted.

To potential side effects and possible resource scarcities, we can add a third category of unintended consequences—*massive disruptions* that have exposed the unprecedented degree of interconnectedness and dependence of our large, human-made systems. Consider three examples: the terrorist attacks against the United States on September 11, 2001; the massive power outage in August 2003 that affected some 10 million people in Ontario and 45 million people in eight U.S. states, known as the “Northeast Blackout of 2003”; and Hurricane Katrina in 2005. As these three events plainly illustrate, with cascading failures, panic, and lack of understanding of emergent behaviors and unexpected consequences, we have entered into a new regime dominated by technology that is often not in harmony with underlying systems, social norms, and regulatory structures.

In parallel, human civilization’s impact on the natural environment—increasingly visible, measurable, and consequential—continues to grow and expand, from overfishing of the oceans to the possible melting of the polar ice caps to changes in soil due to large-scale monoagriculture (growing of a single crop year after year over a wide area). In each case, the interaction of systems is vital.

The metamorphosis to systems that interact more and more—*systems of systems*—continues to accelerate. Our new century is characterized by continued population growth and an ongoing transformation of how we communicate globally, both with others and with the physical world, thanks to the Internet. The separation between the three distinct systems we introduced earlier—transportation “embodied” in the automobile, communications “embodied” in the telephone, and energy “embodied” in the light bulb—continues to shrink (see figure 1.2). To be sure, the systems remain distinct—for example, the highway system and the Internet have not merged into one supersystem. Yet, they are linked to a far greater degree, as evidenced by the transponders in cars used to track vehicles and charge tolls through an Internet-based payment process connecting the automobile to the banking system.

Consider the Global Positioning System (GPS). In 1972, the U.S. Air Force tested two prototype GPS receivers it wanted to use for precision navigation of intercontinental bombers. Six years later, the first experimental GPS satellite was launched. And in 1983, President Reagan—responding to the Soviet downing of Korea Air Lines flight 007—issued a directive that GPS be made freely available for nonmilitary use. It has taken only 25 years for GPS to become an option for every outdoorsman or new car buyer, and it is the rare taxi that doesn’t have a GPS available—for better or worse.

The transportation system and the communications system are linked via the GPS. And the technology has much further to go. Visionary engineers have been developing Intelligent Transportation Systems (ITS) to minimize congestion and enhance efficiency and safety. GPS technology

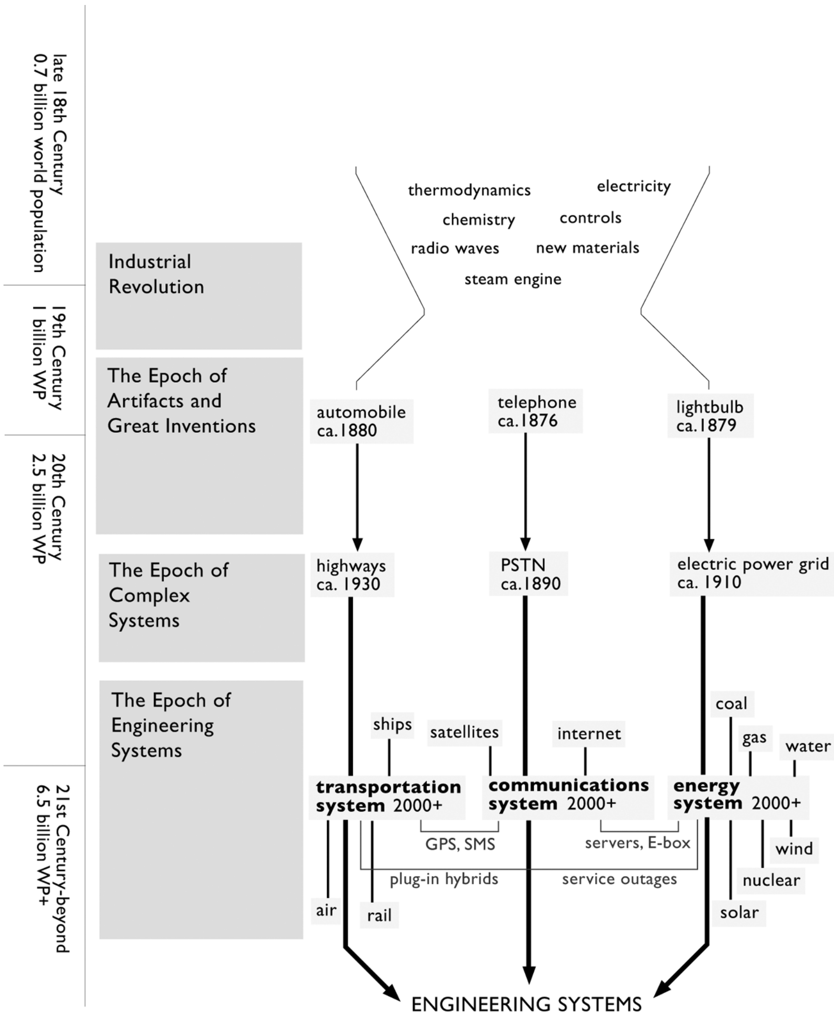

Figure 1.2
Major epochs in the evolution toward engineering systems. The transportation, communications, and energy systems are the major "spines" of this book and show significantly higher levels of complexity and integration over time (eighteenth century to twenty-first century).

is already standard in higher-end cars and is at the center of the next generation of smart cars that will guide us around traffic jams effortlessly. And those vehicles will likely be plug-in hybrids that use electricity, not only gasoline—thus hitching the transportation and communications systems to the energy system on a massive scale. The result? A *mobility* system—bigger and more complex, but also full of new opportunities and new challenges.

What happens if widespread adoption of hybrid or fully electric cars achieves desired reductions in petroleum consumption, but dramatically increases the challenge to the electrical grid of reliably supplying power to current users? Some believe the millions of car batteries in homeowners' garages could be used as energy storage devices, with the potential to smooth spikes in peak electricity demand; others believe these battery systems would be too valuable to use in such a way. This illustrates that there are both opportunities *and* challenges—and they are not all strictly technical. Some are decidedly in the sociopolitical realm. This evolution and linking of technological artifacts, enabling networks, the natural environment, and human agents is the domain of engineering systems, highly technical and socially complex systems that aim to fulfill important functions in society.

What is perhaps most notable about today's engineering systems is that the ways in which they challenge us stem, in large part, directly from the artifacts that began their evolution. For instance, building on earlier technologies, scientists and engineers have brought us major advances in health care such as surgical techniques that sometimes push the boundaries of science fiction. Advances have led to a health care system that contributes to increased longevity and quality of life, but at the same time consumes increasingly large portions of the national and individual budgets. Does the answer to the problem lie exclusively in technology, in a new artifact? We do not think so, and argue throughout this book that what is called for is a combined social and technological approach. For health care, there is no easy answer, but superb technology coupled with superb policy design may represent the correct solution, even if it is difficult to determine.

From Engineering to Engineering Systems

What makes these challenges so difficult to meet is precisely the degree to which they go beyond what might be called the "old" definition of engineering, which focuses on the technical aspect of the work: creating a needed *something* (an artifact) that functions safely and at minimal cost. It is a definition that fits well with inventors who were also engineers, such as Henry Ford and Thomas Edison.

An engineer of old may have been called on to address a relatively narrow problem, say, increasing the throughput of a given machine in a factory or reducing the probability that an electrical component will short out in a motor. He might have been asked to design every aspect of a chemical plant, from the equipment to the processes—a larger challenge, but still relatively narrow. He may even have been assigned to figure out the technology needed to put a man on the moon or assemble and operate an international space station in Earth orbit—tremendous accomplishments, and some of the most complex systems humans have ever created.

Today, working in an engineering system, that same engineer has to interact with a host of socioeconomic complexities and "externalities"—impacts, either positive or negative, that are not a direct part of the artifact or even a self-contained system or process under consideration. It used to be that engineers, even those who were beginning to understand that these externalities might matter, did not worry about them in their designs. Today, these externalities must be factored into the design process. It is all about broadening the boundaries of the system, because the system has, in some ways, swamped us. No one cared about auto emissions until there were so many cars that the emissions began to choke our world. No one cared about how cars were fueled until it became apparent that extracting and refining oil was becoming increasingly difficult and expensive. Unintended consequences and system interactions are becoming the norm, not the exception. And the challenges are even greater, because the externalities are more complex than we ever imagined.

The world has truly changed for the engineer. Speaking of the "engineer as problem solver," Dr. Subra Suresh—when Dean of the MIT School of Engineering—noted that the great accomplishments of the eighteenth through early twentieth centuries nevertheless "created their own set of shortfalls or negative impacts on society." He described how the accomplishments of the twentieth century "brought social and technical changes on a broad scale—but engineering did not generally include social sciences and long-term societal impact." Most of the great challenges engineers face in the twenty-first century, he explained, involve fixing the successes of the greatest achievements of the twentieth century.[14]

Again, consider the transportation system, and particularly the automobile. It is a significant technical achievement that has delivered a tremendous amount of societal value and individual freedom, including personal mobility, but is also fraught with unintended consequences. In the beginning of its development, there was an open competition for which design would prevail. Steam-powered and electric-powered vehicles competed for decades with the gasoline-powered vehicle and its internal combustion engine, which became the dominant design in the 1910s (largely because of issues of energy storage problems for electric cars that are still unresolved). Settling that competition, along with advances in manufacturing, made it possible for millions of automobiles to be made and sold.

None of the early developers of the automobile were out to create traffic jams, but that's part of what their artifacts have wrought. After all, these millions of cars needed roads to drive on. Despite the best of intentions, no engineer could figure out exactly how many roads, where they ought to go, or what types of roads were needed for the future— at least not with a high degree of certainty. A network of highways and roads had to evolve, from the streets in our neighborhoods to Route 66 and from the nearest interstate to Boston's Big Dig (we discuss the latter in some detail in chapter 6).

With so much congestion on these roads, one of the first solutions posed was always to build more highways, or widen the existing ones. That may decrease the effect of one externality in a positive way but increase the negative effects of another externality. After all, if the roads are clearer, might more people drive? And if they do, might that not increase pollution? Also, how much of our land do we want to dedicate to transport rather than, say, fields or woods?

Another example is the telephone—the initial artifact in today's modern communication system. At first blush, it's hard to find anything bad about the telephone—except perhaps the calls from telemarketers that always seem to come just as you sit down for dinner. But dig deeper, and the externalities begin to emerge. For example, the advent of the telephone is widely considered as the beginning of the decline in letter writing. Moreover, the telephone brought us the cell phone. Kids are constantly texting on their cell phones.[15] People are talking on these phones while driving, and even texting from behind the wheel. That last behavior is a very dangerous proposition; a study by the Virginia Tech Transportation Institute found that long-haul truck drivers who were texting while driving were 23 times more likely to be involved in a traffic accident than drivers who were not texting.[16] Now imagine those trucks involved in accidents with a new generation of lighter-weight cars that may be less safe!

It seems that everything an engineer must deal with has changed—tremendously—requiring a much broader perspective. The "big" ships of old have grown to such a size over the decades that everything about them, from how they are powered to the ports where they dock to their effects on those ports, has had to be reconsidered. Even the Panama Canal is being expanded to accommodate the new fleet of supertankers; some figures suggest that more than a third of the world's container ships will be too large for the "old" canal by the end of 2011. Today's automotive engine is almost 100 times more powerful per pound than a century ago, creating a host of more or less obvious ramifications. Each and every artifact and the system it has spawned has become a legacy, affecting millions of users, potentially, with every small adjustment in that system. Figure 1.3 shows some of the tremendous progress humans have made in technical capabilities.[17] Recent work has shown clearly that large-scale exponential change and improvements—like those first noted by Intel cofounder Gordon E. Moore for semiconductors—are pervasive throughout technology.[18] These technological improvements suggest strongly that viewing technology as only a source of problems is short-sighted; rather, through exponential improvements technology is also part of the potential future solutions to problems.

Some have said that engineers are out to fix the world. Engineers might say they have a great solution to a given problem, but politicians are getting in the way. In engineering systems, we seek ways to codesign the artifacts and every other element of the system, along with the political system and regulations, and in the context of societal attitudes and norms.

The discipline of engineering systems deals not only with the kinds of increases in scale, scope, and complexity you've read about in this chapter, but also with changes in technology that continue to amaze those of us who can remember the days before, say, personal computers. Add to this the changes in how decisions get made, the growing demand for life-cycle thinking, and the continual shift in what society values about the world and how it works, and you begin to see how much is at stake.

And that doesn't even touch on half of what has changed for the engineer. Designs need to address quality in new ways, and find "perfection" from the beginning. Systems need to be flexible where once they could be much more rigid. Everything moves faster, and the rate of

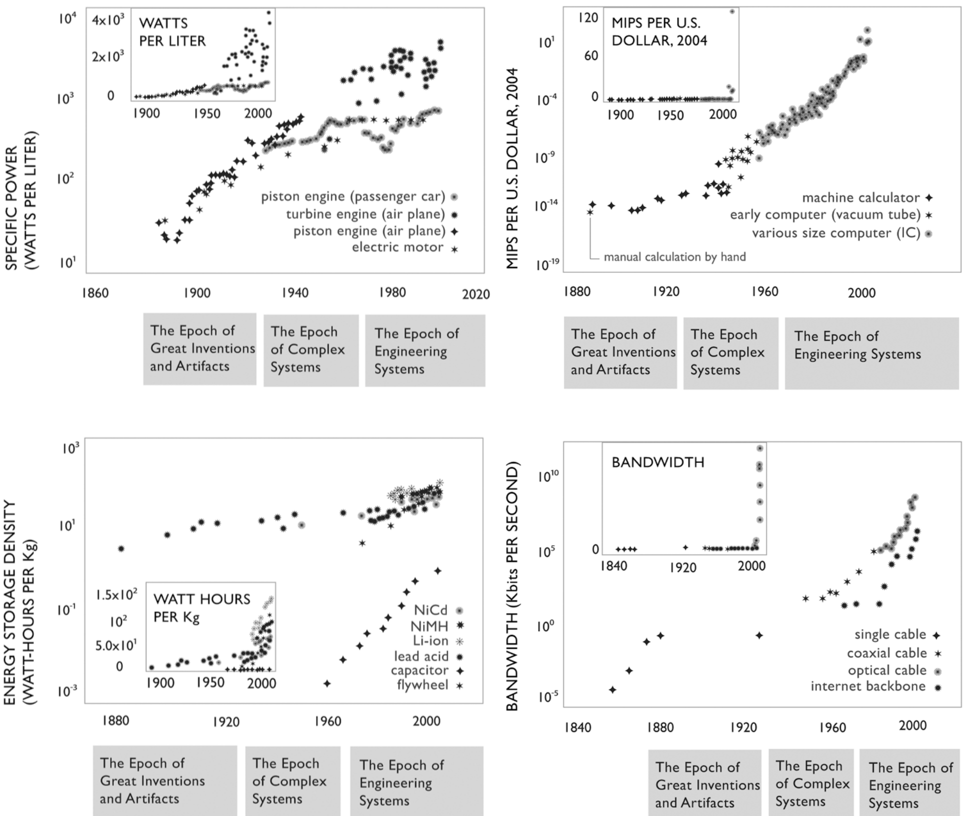

Figure 1.3
Progress of technological capabilities (1860–2010): specific power (upper left), energy storage density (lower left), normalized cost of computational speed (upper right), and communication bandwidth (lower right). Data are from Koh and Magee (2006) and Koh and Magee (2008), respectively.

technological and social change shows no signs of slowing down. Stakeholders who were never on the old radar screen must today be given the opportunity to have real input into solutions that engineers may help develop. Assumptions of the past give way to new realities that could not have been imagined even a generation ago—like all the new shipping lanes across the Arctic that are opening up as the polar ice cap gradually disappears.

All of this plays out in an increasingly global context within which engineers must work. One of our MIT colleagues, Noelle Eckley Selin,[19] provides a superb example: the problem of mercury emissions that are poisoning our environment, which she has been studying for some time.

Noelle explains that despite political attention in the United States going back at least to the 1950s, and international policymaking since the 1970s, the problem has yet to be solved. Mercury deposition continues to pose risks to human health worldwide, with the pollutant emitted from sources such as coal-fired power plants or different types of industrial and mining activities. Long-lived, elemental mercury then circulates globally in the atmosphere, only to fall from the sky and accumulate in fish as toxic methylmercury.

It turns out that mercury emissions can be reduced with improvements in emissions control technologies—so there is at least a partial, old-style engineering solution. But even if those controls were implemented in the United States, it wouldn't eliminate the poison. This is where the evolution to an engineering systems point of view makes the head spin, as the engineer tries to find a solution given the global circulation of the mercury.

Here's the specific issue. While Florida has the highest mercury deposition in the United States, less than 20 percent of this deposition comes from domestic sources, most of which are in the industrial Midwest. The mercury from the Midwest can either deposit locally in the region or take an atmospheric ride around the world, combining with deposits from countries such as China, before finding its way into Florida's waterways.[20] For this reason, no solution based on technology alone will do the trick, because the effectiveness of its implementation will depend on the resolution of the global political issues associated with mercury in the atmosphere (see figure 1.4).

As Noelle told us, "Effectively addressing mercury exposure requires taking into account the system dynamics and complexities of the mercury problem at various spatial and temporal scales, including a global perspective." In other words, a technical approach is essential even to understanding the problem, but alone it can get us only so far. It is an engineering systems problem that must be seen as sociotechnical and that demands solutions rooted not only in technology but also in the social sciences and management. The Chinese government has to be part of the mercury equation. The ultimate answer may differ by region, country, or even continent. That is the reality of a world in which, for example, some countries rely entirely on nuclear and hydroelectric power for their energy needs, while others still rely heavily on fossil fuels.

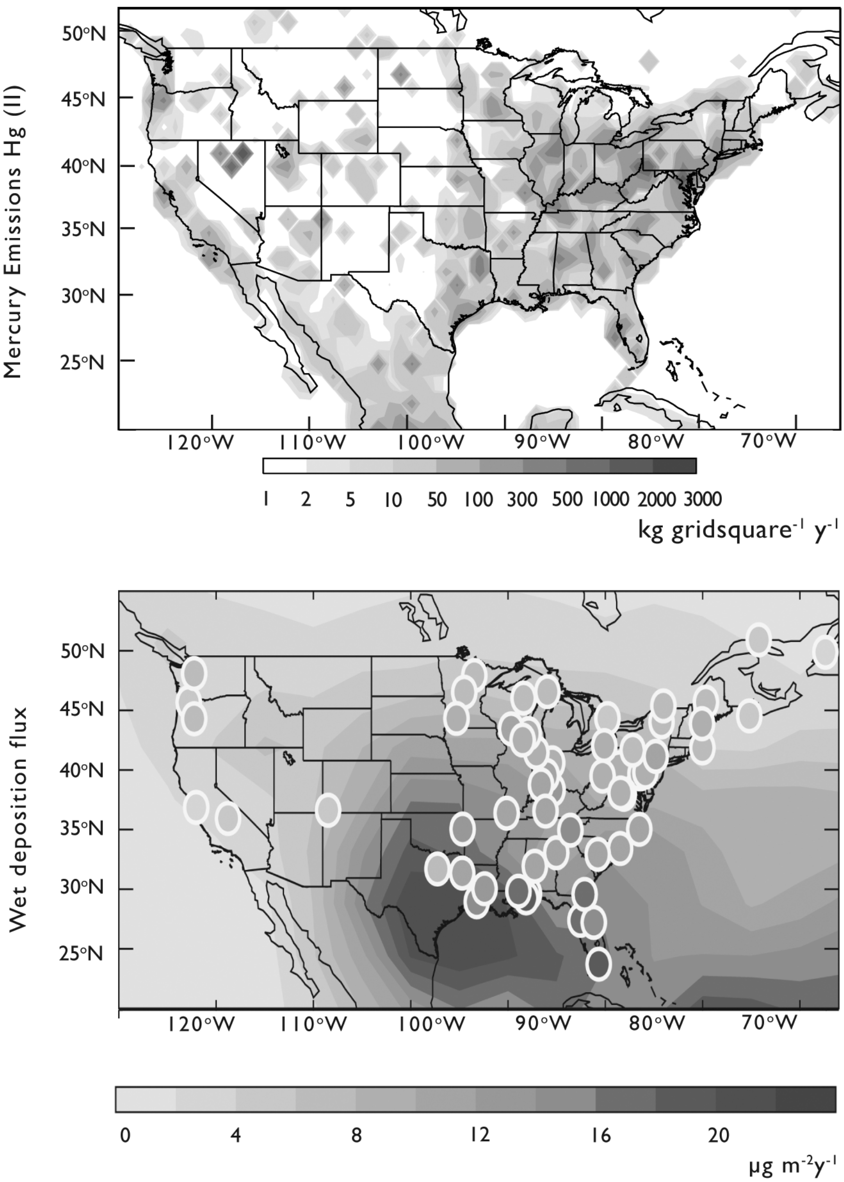

Figure 1.4
Top panel: Annual anthropogenic emissions of Hg(II), the form of mercury that deposits on a regional scale (kilograms per year per one degree by one degree grid square) in North America for 2000. Data are from Pacyna et al. (2006). Bottom panel: Annual mean wet deposition flux of mercury over the United States for 2004–2005 (micrograms per square meter per year). Observations from 57 sites of the Mercury Deposition Network (MDN) shown as ovals are compared to the GEOS-Chem model results (background).

Indeed, impacts are global, but standards, cultural preferences, and other factors are not always the same. A software company such as Google that wants to be global, for instance, faces the daunting challenge of figuring out how to deal with different and often conflicting regulatory regimes of different countries in the global markets, which play an important role in shaping each national market. These national differences produce large variations in the nature of modern systems throughout the world.

The engineering systems we have described in brief here—transportation, energy, and communications—along with others that are essential such as modern medicine and health care, are what have enabled us as humans to transform our economies from agrarian to industrial and bring us into the information and service age. They are what allow us to be a global society.

In the next chapter, we take a closer look at how engineering systems are defined, characterized, and classified in terms of their functions, inherent structure, and sociotechnical complexity.

Suggestions for Supplemental Reading

David H. Bain (1999), *Empire Express: Building the First Transcontinental Railroad*, New York: Viking.

Rachel Carson (1962), *Silent Spring*, Boston, MA: Houghton Mifflin.

Robert William Fogel (2004), *The Escape from Hunger and Premature Death, 1700–2100: Europe, America, and the Third World*, Cambridge, England: Cambridge University Press.

Thomas P. Hughes (1983), *Networks of Power: Electrification in Western Society, 1880–1930*, Baltimore, MD: Johns Hopkins University Press.

Donella H. Meadows, Dennis L. Meadows, Jorgen Randers, and William W. Behrens (1972), *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament for Mankind*, New York: Universe Books.

Joel Mokyr (2010), *The Enlightened Economy: An Economic History of Britain 1700–1850*, New Haven, CT: Yale University Press.

Kenneth Pomeranz (2000), *The Great Divergence: China, Europe, and the Making of the Modern World Economy*, Princeton, NJ: Princeton University Press.