CHAPTER ONE

The Six Epochs

> Everyone takes the limits of his own vision for the limits of the world.
>
> —ARTHUR SCHOPENHAUER

I am not sure when I first became aware of the Singularity. I'd have to say it was a progressive awakening. In the almost half century that I've immersed myself in computer and related technologies, I've sought to understand the meaning and purpose of the continual upheaval that I have witnessed at many levels. Gradually, I've become aware of a transforming event looming in the first half of the twenty-first century. Just as a black hole in space dramatically alters the patterns of matter and energy accelerating toward its event horizon, this impending Singularity in our future is increasingly transforming every institution and aspect of human life, from sexuality to spirituality.

What, then, is the Singularity? It's a future period during which the pace of technological change will be so rapid, its impact so deep, that human life will be irreversibly transformed. Although neither utopian nor dystopian, this epoch will transform the concepts that we rely on to give meaning to our lives, from our business models to the cycle of human life, including death itself. Understanding the Singularity will alter our perspective on the significance of our past and the ramifications for our future. To truly understand it inherently changes one's view of life in general and one's own particular life. I regard someone who understands the Singularity and who has reflected on its implications for his or her own life as a "singularitarian."[1]

I can understand why many observers do not readily embrace the obvious implications of what I have called the law of accelerating returns (the inherent acceleration of the rate of evolution, with technological evolution as a continuation of biological evolution), After all, it took me forty years to be able to see what was right in front of me, and I still cannot say that I am entirely comfortable with all of its consequences.

The key idea underlying the impending Singularity is that the pace of change of our human-created technology is accelerating and its powers are expanding at an exponential pace. Exponential growth is deceptive. It starts out almost imperceptibly and then explodes with unexpected fury—unexpected, that is, if one does not take care to follow its trajectory. (See the "Linear vs. Exponential Growth" graph on p. 10.)

Consider this parable: a lake owner wants to stay at home to tend to the lake's fish and make certain that the lake itself will not become covered with lily pads, which are said to double their number every few days. Month after month, he patiently waits, yet only tiny patches of lily pads can be discerned, and they don't seem to be expanding in any noticeable way. With the lily pads covering less than 1 percent of the lake, the owner figures that it's safe to take a vacation and leaves with his family. When he returns a few weeks later, he's shocked to discover that the entire lake has become covered with the pads, and his fish have perished. By doubling their number every few days, the last seven doublings were sufficient to extend the pads' coverage to the entire lake. (Seven doublings extended their reach 128-fold.) This is the nature of exponential growth.

Consider Gary Kasparov, who scorned the pathetic state of computer chess in 1992. Yet the relentless doubling of computer power every year enabled a computer to defeat him only five years later.[2] The list of ways computers can now exceed human capabilities is rapidly growing. Moreover, the once narrow applications of computer intelligence are gradually broadening in one type of activity after another. For example, computers are diagnosing electrocardiograms and medical images, flying and landing airplanes, controlling the tactical decisions of automated weapons, making credit and financial decisions, and being given responsibility for many other tasks that used to require human intelligence. The performance of these systems is increasingly based on integrating multiple types of artificial intelligence (AI). But as long as there is an AI shortcoming in any such area of endeavor, skeptics will point to that area as an inherent bastion of permanent human superiority over the capabilities of our own creations.

This book will argue, however, that within several decades information-based technologies will encompass all human knowledge and proficiency, ultimately including the pattern-recognition powers, problem-solving skills, and emotional and moral intelligence of the human brain itself.

Although impressive in many respects, the brain suffers from severe limitations. We use its massive parallelism (one hundred trillion interneuronal connections operating simultaneously) to quickly recognize subtle patterns. But our thinking is extremely slow: the basic neural transactions are several million times slower than contemporary electronic circuits. That makes our physiological bandwidth for processing new information extremely limited compared to the exponential growth of the overall human knowledge base.

Our version 1.0 biological bodies are likewise frail and subject to a myriad of failure modes, not to mention the cumbersome maintenance rituals they require. While human intelligence is sometimes capable of soaring in its creativity and expressiveness, much human thought is derivative, petty, and circumscribed.

The Singularity will allow us to transcend these limitations of our biological bodies and brains. We will gain power over our fates. Our mortality will be in our own hands. We will be able to live as long as we want (a subtly different statement from saying we will live forever). We will fully understand human thinking and will vastly extend and expand its reach. By the end of this century, the nonbiological portion of our intelligence will be trillions of trillions of times more powerful than unaided human intelligence.

We are now in the early stages of this transition. The acceleration of paradigm shift (the rate at which we change fundamental technical approaches) as well as the exponential growth of the capacity of information technology are both beginning to reach the "knee of the curve," which is the stage at which an exponential trend becomes noticeable. Shortly after this stage, the trend quickly becomes explosive. Before the middle of this century, the growth rates of our technology—which will be indistinguishable from ourselves—will be so steep as to appear essentially vertical. From a strictly mathematical perspective, the growth rates will still be finite but so extreme that the changes they bring about will appear to rupture the fabric of human history. That, at least, will be the perspective of unenhanced biological humanity.

The Singularity will represent the culmination of the merger of our biological thinking and existence with our technology, resulting in a world that is still human but that transcends our biological roots. There will be no distinction, post-Singularity, between human and machine or between physical and virtual reality. If you wonder what will remain unequivocally human in such a world, it's simply this quality: ours is the species that inherently seeks to extend its physical and mental reach beyond current limitations.

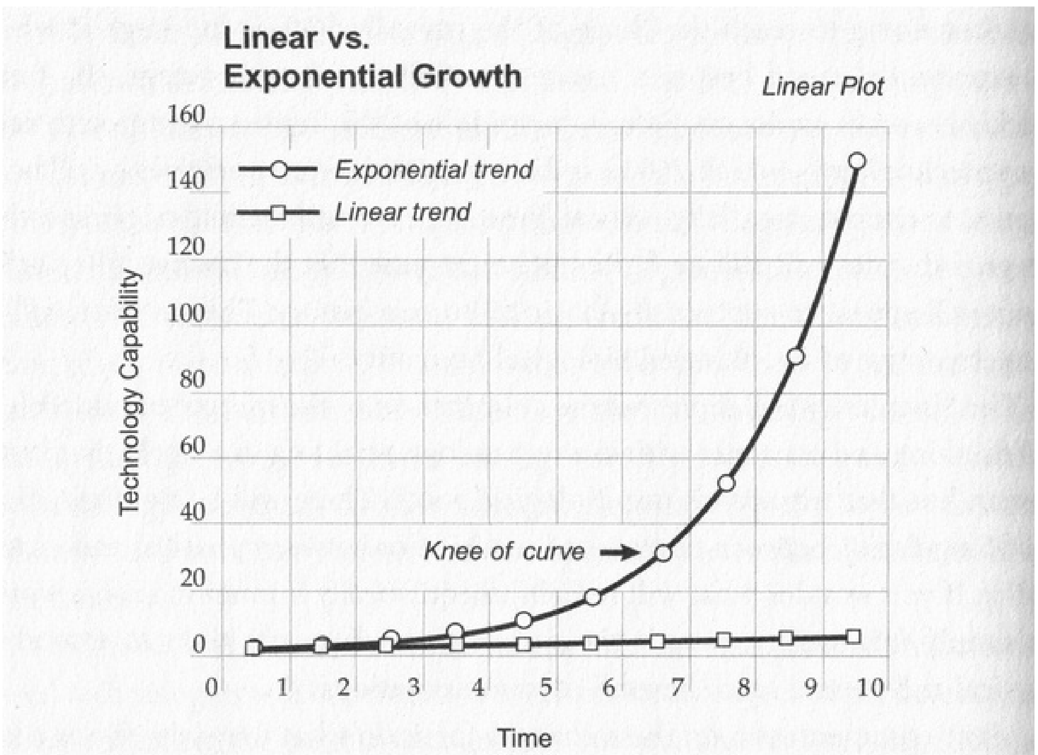
Many commentators on these changes focus on what they perceive as a loss of some vital aspect of our humanity that will result from this transition. This perspective stems, however, from a misunderstanding of what our technology will become. All the machines we have met to date lack the essential subtlety of human biological qualities. Although the Singularity has many faces, its most important implication is this: our technology will match and then vastly exceed the refinement and suppleness of what we regard as the best of human traits.

The Intuitive Linear View Versus the Historical Exponential View

> When the first transhuman intelligence is created and launches itself into recursive self-improvement, a fundamental discontinuity is likely to occur, the likes of which I can't even begin to predict.
>
> —MICHAEL ANISSIMOV

In the 1950s John von Neumann, the legendary information theorist, was quoted as saying that "the ever-accelerating progress of technology ... gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue."[3] Von Neumann makes two important observations here: *acceleration* and *singularity*.

The first idea is that human progress is exponential (that is, it expands by repeatedly multiplying by a constant) rather than linear (that is, expanding by repeatedly adding a constant).



Linear versus exponential: **Linear growth is steady; exponential growth becomes explosive.**

The second is that exponential growth is seductive, starting out slowly and virtually unnoticeably, but beyond the knee of the curve it turns explosive and profoundly transformative. The future is widely misunderstood. Our forebears expected it to be pretty much like their present, which had been pretty much like their past. Exponential trends did exist one thousand years ago, but they were at that very early stage in which they were so flat and so slow that they looked like no trend at all. As a result, observers' expectation of an unchanged future was fulfilled. Today, we anticipate continuous technological progress and the social repercussions that follow. But the future will be far more surprising than most people realize, because few observers have truly internalized the implications of the fact that the rate of change itself is accelerating.

Most long-range forecasts of what is technically feasible in future time periods dramatically underestimate the power of future developments because they are based on what I call the "intuitive linear" view of history rather than the "historical exponential" view. My models show that we are doubling the paradigm-shift rate every decade, as I will discuss in the next chapter. Thus the twentieth century was gradually speeding up to today's rate of progress; its achievements, therefore, were equivalent to about twenty years of progress at the rate in 2000. We'll make another twenty years of progress in just fourteen years (by 2014), and then do the same again in only seven years. To express this another way, we won't experience one hundred years of technological advance in the twenty-first century; we will witness on the order of twenty thousand years of progress (again, when measured by *today's* rate of progress), or about one thousand times greater than what was achieved in the twentieth century.[4]

Misperceptions about the shape of the future come up frequently and in a variety of contexts. As one example of many, in a recent debate in which I took part concerning the feasibility of molecular manufacturing, a Nobel Prizewinning panelist dismissed safety concerns regarding nanotechnology, proclaiming that "we're not going to see self-replicating nanoengineered entities [devices constructed molecular fragment by fragment] for a hundred years." I pointed out that one hundred years was a reasonable estimate and actually matched my own appraisal of the amount of technical progress required to achieve this particular milestone when measured at *today's rate of progress* (five times the average rate of change we saw in the twentieth century). But because we're doubling the rate of progress every decade, we'll see the equivalent of a century of progress—*at today's rate*—in only twenty-five calendar years.

Similarly at *Time* magazine's Future of Life conference, held in 2003 to celebrate the fiftieth anniversary of the discovery of the structure of DNA, all of the invited speakers were asked what they thought the next fifty years would be like.[5] Virtually every presenter looked at the progress of the last fifty years and used it as a model for the next fifty years. For example, James Watson, the codiscoverer of DNA, said that in fifty years we will have drugs that will allow us to eat as much as we want without gaining weight.

I replied, "Fifty years?" We have accomplished this already in mice by blocking the fat insulin receptor gene that controls the storage of fat in the fat cells. Drugs for human use (using RNA interference and other techniques we will discuss in chapter 5) are in development now and will be in FDA tests in several years. These will be available in five to ten years, not fifty. Other projections were equally shortsighted, reflecting contemporary research priorities rather than the profound changes that the next half century will bring. Of all the thinkers at this conference, it was primarily Bill Joy and I who took account of the exponential nature of the future, although Joy and I disagree on the import of these changes, as I will discuss in chapter 8.

People intuitively assume that the current rate of progress will continue for future periods. Even for those who have been around long enough to experience how the pace of change increases over time, unexamined intuition leaves one with the impression that change occurs at the same rate that we have experienced most recently. From the mathematician's perspective, the reason for this is that an exponential curve looks like a straight line when examined for only a brief duration. As a result, even sophisticated commentators, when considering the future, typically extrapolate the current pace of change over the next ten years or one hundred years to determine their expectations. This is why I describe this way of looking at the future as the "intuitive linear" view.

But a serious assessment of the history of technology reveals that technological change is exponential. Exponential growth is a feature of any evolutionary process, of which technology is a primary example. You can examine the data in different ways, on different timescales, and for a wide variety of technologies, ranging from electronic to biological, as well as for their implications, ranging from the amount of human knowledge to the size of the economy. The acceleration of progress and growth applies to each of them. Indeed, we often find not just simple exponential growth, but "double" exponential growth, meaning that the rate of exponential growth (that is, the exponent) is itself growing exponentially (for example, see the discussion on the price-performance of computing in the next chapter).

Many scientists and engineers have what I call "scientist's pessimism." Often, they are so immersed in the difficulties and intricate details of a contemporary challenge that they fail to appreciate the ultimate long-term implications of their own work, and the larger field of work in which they operate. They likewise fail to account for the far more powerful tools they will have available with each new generation of technology.

Scientists are trained to be skeptical, to speak cautiously of current research goals, and to rarely speculate beyond the current generation of scientific pursuit. This may have been a satisfactory approach when a generation of science and technology lasted longer than a human generation, but it does not serve society's interests now that a generation of scientific and technological progress comprises only a few years.

Consider the biochemists who, in 1990, were skeptical of the goal of transcribing the entire human genome in a mere fifteen years. These scientists had just spent an entire year transcribing a mere one ten-thousandth of the genome. So, even with reasonable anticipated advances, it seemed natural to them that it would take a century, if not longer, before the entire genome could be sequenced.

Or consider the skepticism expressed in the mid-1980s that the Internet would ever be a significant phenomenon, given that it then included only tens of thousands of nodes (also known as servers). In fact, the number of nodes was doubling every year, so that there were likely to be tens of millions of nodes ten years later. But this trend was not appreciated by those who struggled with state-of-the-art technology in 1985, which permitted adding only a few thousand nodes throughout the world in a single year."[6]

The converse conceptual error occurs when certain exponential phenomena are first recognized and are applied in an overly aggressive manner without modeling the appropriate pace of growth. While exponential growth gains speed over time, it is not instantaneous. The run-up in capital values (that is, stock market prices) during the "Internet bubble" and related telecommunications bubble (1997–2000) was greatly in excess of any reasonable expectation of even exponential growth. As I demonstrate in the next chapter, the actual adoption of the Internet and e-commerce did show smooth exponential growth through both boom and bust; the overzealous expectation of growth affected only capital (stock) valuations. We have seen comparable mistakes during earlier paradigm shifts—for example, during the early railroad era (1830s), when the equivalent of the Internet boom and bust led to a frenzy of railroad expansion.

Another error that prognosticators make is to consider the transformations that will result from a single trend in to day's world as if nothing else will change. A good example is the concern that radical life extension will result in overpopulation and the exhaustion of limited material resources to sustain human life, which ignores comparably radical wealth creation from nanotechnology and strong AI. For example, nanotechnology-based manufacturing devices in the 2020s will be capable of creating almost any physical product from inexpensive raw materials and information.

I emphasize the exponential-versus-linear perspective because it's the most important failure that prognosticators make in considering future trends. Most technology forecasts and forecasters ignore altogether this historical exponential view of technological progress. Indeed, almost everyone I meet has a linear view of the future. That's why people tend to overestimate what can be achieved in the short term (because we tend to leave out necessary details) but underestimate what can be achieved in the long term (because exponential growth is ignored).

The Six Epochs

First we build the tools, then they build us.

—MARSHALL MCLUHAN

The future ain't what it used to be.

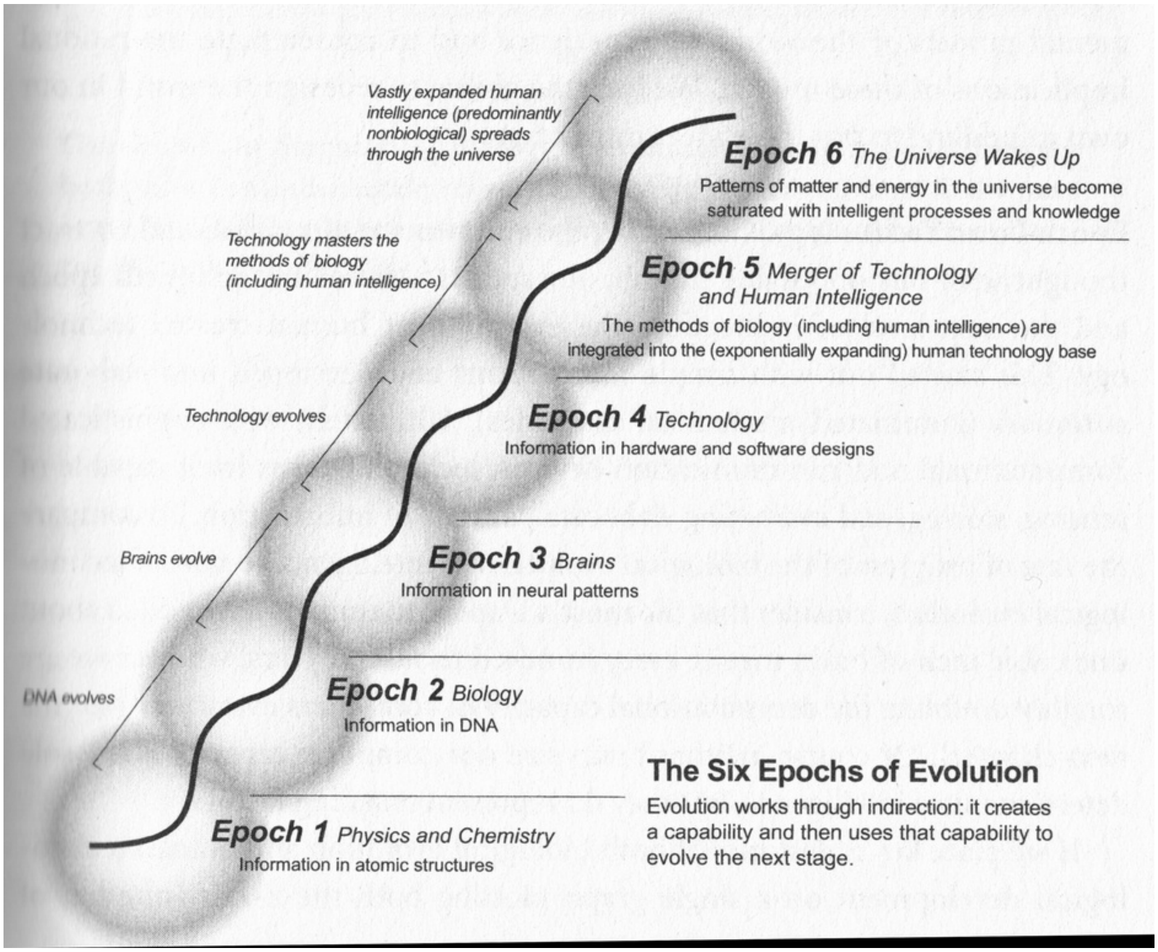
—YOGI BERRA

Evolution is a process of creating patterns of increasing order. I'll discuss the concept of order in the next chapter; the emphasis in this section is on the concept of patterns. I believe that it's the evolution of patterns that constitutes the ultimate story of our world. Evolution works through indirection: each stage or epoch uses the information-processing methods of the previous epoch to create the next. I conceptualize the history of evolution—both biological and technological—as occurring in six epochs. As we will discuss, the Singularity will begin with Epoch Five and will spread from Earth to the rest of the universe in Epoch Six.

Epoch One: Physics and Chemistry. We can trace our origins to a state that represents information in its basic structures: patterns of matter and energy. Recent theories of quantum gravity hold that time and space are broken down into discrete quanta, essentially fragments of information. There is controversy as to whether matter and energy are ultimately digital or analog in nature, but regardless of the resolution of this issue, we do know that atomic structures store and represent discrete information.

A few hundred thousand years after the Big Bang, atoms began to form, as electrons became trapped in orbits around nuclei consisting of protons and neutrons. The electrical structure of atoms made them "sticky." Chemistry was born a few million years later as atoms came together to create relatively stable structures called molecules. Of all the elements, carbon proved to be the most versatile; it's able to form bonds in four directions (versus one to three for most other elements), giving rise to complicated, information-rich, three-dimensional structures.

The rules of our universe and the balance of the physical constants that govern the interaction of basic forces are so exquisitely, delicately, and exactly appropriate for the codification and evolution of information (resulting in increasing complexity) that one wonders how such an extraordinarily unlikely situation came about. Where some see a divine hand, others see our own hands—namely, the anthropic principle, which holds that only in a universe that allowed our own evolution would we be here to ask such questions.[7] Recent theories of physics concerning multiple universes speculate that new universes are created on a regular basis, each with its own unique rules, but that most of these either die out quickly or else continue without the evolution of any interesting patterns (such as Earth-based biology has created) because their rules do not support the evolution of increasingly complex forms.[8] It's hard to imagine how we could test these theories of evolution applied to early cosmology, but it's clear that the physical laws of our universe are precisely what they need to be to allow for the evolution of increasing levels of order and complexity.[9]

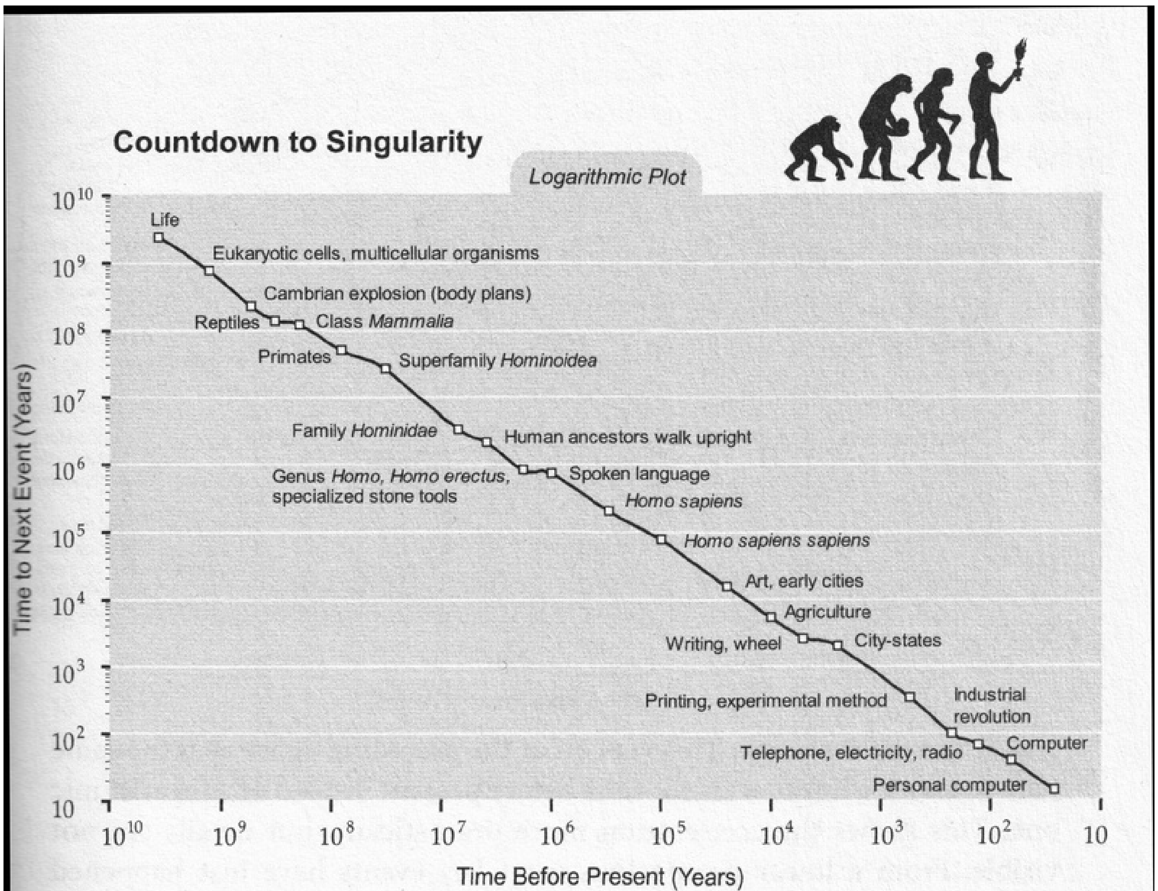


Epoch Two: Biology and DNA. In the second epoch, starting several billion years ago, carbon-based compounds became more and more intricate until complex aggregations of molecules formed self-replicating mechanisms, and life originated. Ultimately, biological systems evolved a precise digital mechanism (DNA) to store information describing a larger society of molecules. This molecule and its supporting machinery of codons and ribosomes enabled a record to be kept of the evolutionary experiments of this second epoch.

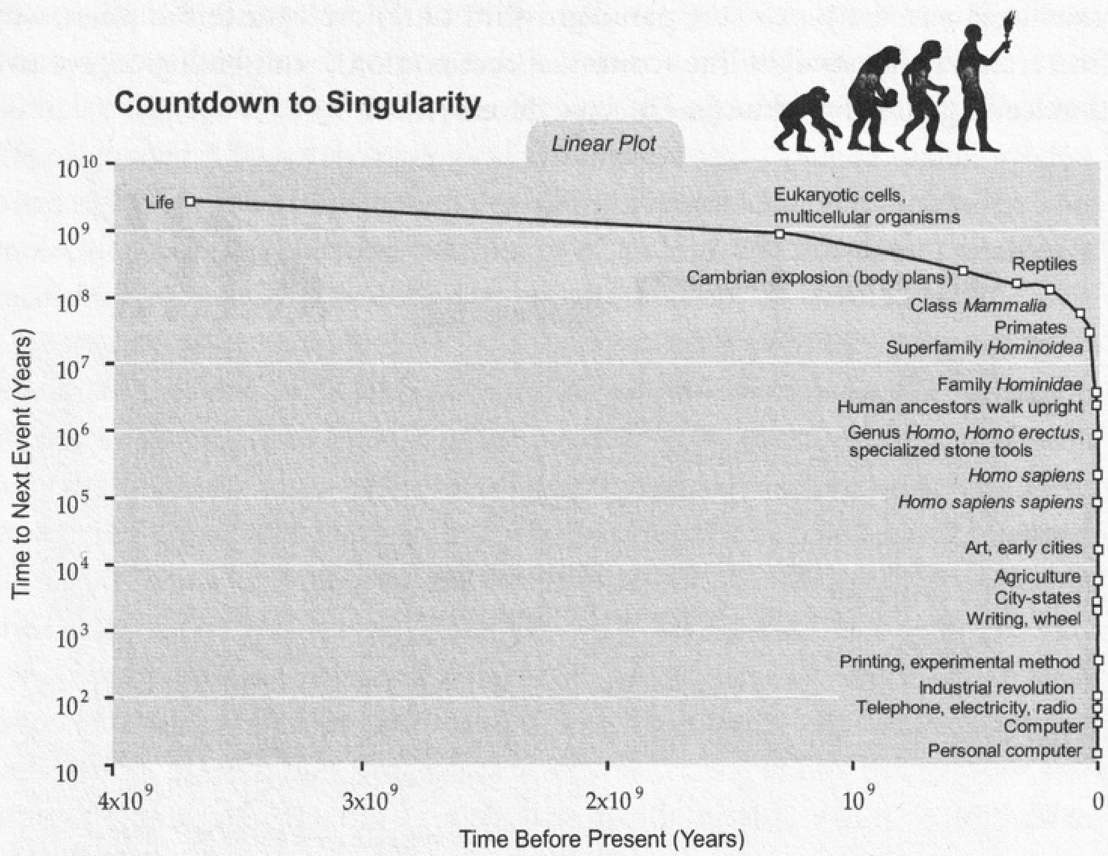
Epoch Three: Brains. Each epoch continues the evolution of information through a paradigm shift to a further level of "indirection." (That is, evolution uses the results of one epoch to create the next.) For example, in the third epoch, DNA-guided evolution produced organisms that could detect information with their own sensory organs and process and store that information in their own brains and nervous systems. These were made possible by second-epoch mechanisms (DNA and epigenetic information of proteins and RNA fragments that control gene expression), which (indirectly) enabled and defined third-epoch information-processing mechanisms (the brains and nervous systems of organisms). The third epoch started with the ability of early animals to recognize patterns, which still accounts for the vast majority of the activity in our brains.[10] Ultimately, our own species evolved the ability to create abstract mental models of the world we experience and to contemplate the rational implications of these models. We have the ability to redesign the world in our own minds and to put these ideas into action.

Epoch Four: Technology. Combining the endowment of rational and abstract thought with our opposable thumb, our species ushered in the fourth epoch and the next level of indirection: the evolution of human-created technology. This started out with simple mechanisms and developed into elaborate automata (automated mechanical machines). Ultimately, with sophisticated computational and communication devices, technology was itself capable of sensing, storing, and evaluating elaborate patterns of information. To compare the rate of progress of the biological evolution of intelligence to that of technological evolution, consider that the most advanced mammals have added about one cubic inch of brain matter every hundred thousand years, whereas we are roughly doubling the computational capacity of computers every year (see the next chapter). Of course, neither brain size nor computer capacity is the sole determinant of intelligence, but they do represent enabling factors.

If we place key milestones of both biological evolution and human technological development on a single graph plotting both the *x*-axis (number of years ago) and the *y*-axis (the paradigm-shift time) on logarithmic scales, we find a reasonably straight line (continual acceleration), with biological evolution leading directly to human-directed development.[11]

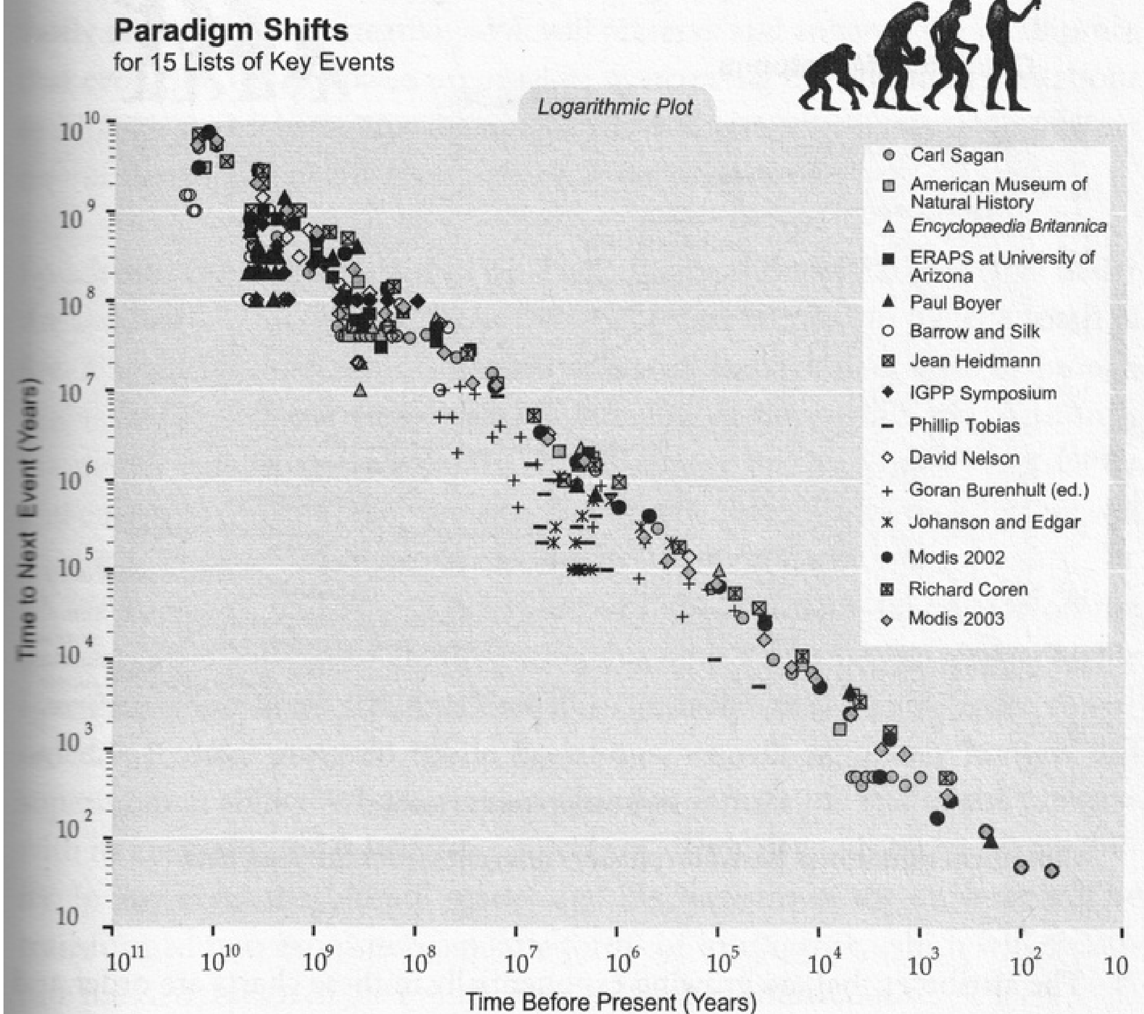


Countdown to Singularity: **Biological evolution and human technology both show continual acceleration, indicated by the shorter time to the next event (two billion years from the origin of life to cells; fourteen years from the PC to the World Wide Web).**



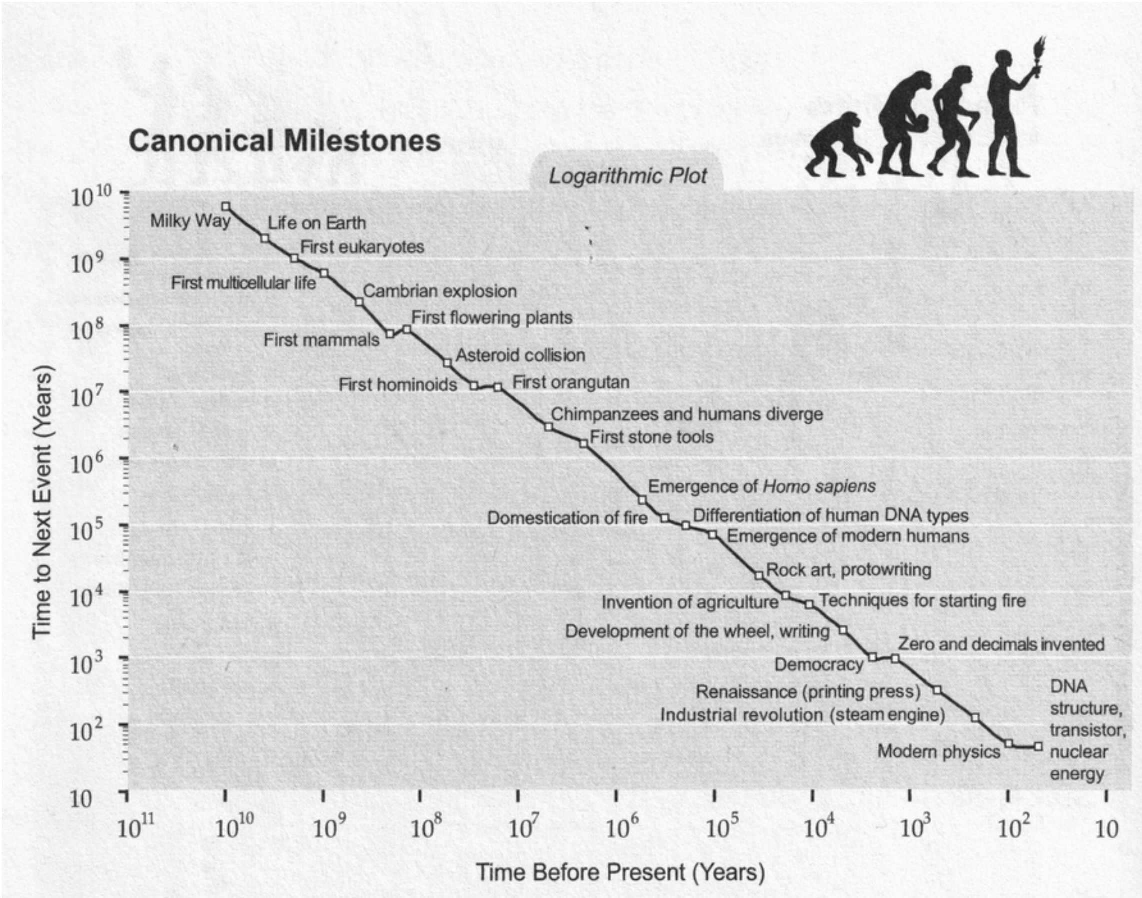
Linear view of evolution: **This version of the preceding figure uses the same data but with a linear scale for time before present instead of a logarithmic one. This shows the acceleration more dramatically, but details are not visible. From a linear perspective, most key events have just happened "recently."**

The above figures reflect my view of key developments in biological and technological history. Note, however, that the straight line, demonstrating the continual acceleration of evolution, does not depend on my particular selection of events. Many observers and reference books have compiled lists of important events in biological and technological evolution, each of which has its own idiosyncrasies. Despite the diversity of approaches, however, if we combine lists from a variety of sources (for example, the *Encyclopaedia Britannica*, the American Museum of Natural History, Carl Sagan's "cosmic calendar," and others), we observe the same obvious smooth acceleration. The following plot combines fifteen different lists of key events.[12] Since different thinkers assign different dates to the same event, and different lists include similar or overlapping events selected according to different criteria, we see an expected "thickening" of the trend line due to the "noisiness" (statistical variance) of this data. The overall trend, however, is very clear.



Fifteen views of evolution: **Major paradigm shifts in the history of the world, as seen by fifteen different lists of key events. There is a clear trend of smooth acceleration through biological and then technological evolution.**

Physicist and complexity theorist Theodore Modis analyzed these lists and determined twenty-eight clusters of events (which he called canonical milestones) by combining identical, similar, and/or related events from the different lists.[13] This process essentially removes the "noise" (for example, the variability of dates between lists) from the lists, revealing again the same progression:



Canonical milestones based on clusters of events from thirteen lists.

The attributes that are growing exponentially in these charts are order and complexity, concepts we will explore in the next chapter. This acceleration matches our commonsense observations. A billion years ago, not much happened over the course of even one million years. But a quarter-million years ago epochal events such as the evolution of our species occurred in time frames of just one hundred thousand years. In technology, if we go back fifty thousand years, not much happened over a one-thousand-year period. But in the recent past, we see new paradigms, such as the World Wide Web, progress from inception to mass adoption (meaning that they are used by a quarter of the population in advanced countries) within only a decade.

Epoch Five: The Merger of Human Technology with Human Intelligence. Looking ahead several decades, the Singularity will begin with the fifth epoch. It will result from the merger of the vast knowledge embedded in our own brains with the vastly greater capacity, speed, and knowledge-sharing ability of our technology. The fifth epoch will enable our human-machine civilization to transcend the human brain's limitations of a mere hundred trillion extremely slow connections.[14]

The Singularity will allow us to overcome age-old human problems and vastly amplify human creativity. We will preserve and enhance the intelligence that evolution has bestowed on us while overcoming the profound limitations of biological evolution. But the Singularity will also amplify the ability to act on our destructive inclinations, so its full story has not yet been written.

Epoch Six: The Universe Wakes Up. I will discuss this topic in chapter 6, under the heading "...on the Intelligent Destiny of the Cosmos." In the aftermath of the Singularity, intelligence, derived from its biological origins in human brains and its technological origins in human ingenuity, will begin to saturate the matter and energy in its midst. It will achieve this by reorganizing matter and energy to provide an optimal level of computation (based on limits we will discuss in chapter 3) to spread out from its origin on Earth.

We currently understand the speed of light as a bounding factor on the transfer of information. Circumventing this limit has to be regarded as highly speculative, but there are hints that this constraint may be able to be superseded.[15] If there are even subtle deviations, we will ultimately harness this superluminal ability. Whether our civilization infuses the rest of the universe with its creativity and intelligence quickly or slowly depends on its immutability. In any event the "dumb" matter and mechanisms of the universe will be transformed into exquisitely sublime forms of intelligence, which will constitute the sixth epoch in the evolution of patterns of information.

This is the ultimate destiny of the Singularity and of the universe.

The Singularity Is Near

> You know, things are going to be really different! ... No, no, I mean really different!
>
> —MARK MILLER (COMPUTER SCIENTIST) TO ERIC DREXLER, AROUND 1986

> What are the consequences of this event? When greater-than-human intelligence drives progress, that progress will be much more rapid. In fact, there seems no reason why progress itself would not involve the creation of still more intelligent entities—on a still-shorter time scale. The best analogy that I see is with the evolutionary past: Animals can adapt to problems and make inventions, but often no faster than natural selection can do its work—the world acts as its own simulator in the case of natural selection. We humans have the ability to internalize the world and conduct "what if's" in our heads; we can solve many problems thousands of times faster than natural selection. Now, by creating the means to execute those simulations at much higher speeds, we are entering a regime as radically different from our human past as we humans are from the lower animals. From the human point of view, this change will be a throwing away of all the previous rules, perhaps in the blink of an eye, an exponential runaway beyond any hope of control.
>
> —VERNOR VINGE, "THE TECHNOLOGICAL SINGULARITY," 1993

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make.
>
> —IRVING JOHN GOOD, "SPECULATIONS CONCERNING THE FIRST ULTRAINTELLIGENT MACHINE," 1965

To put the concept of Singularity into further perspective, let's explore the history of the word itself. "Singularity" is an English word meaning a unique event with, well, singular implications. The word was adopted by mathematicians to denote a value that transcends any finite limitation, such as the explosion of magnitude that results when dividing a constant by a number that gets closer and closer to zero. Consider,