Intelligent & Human Futures
WISDOM ACADEMIC https://doi.org/10.63808/ihf.v2i3.505 Volume 2 , Issue 3

Research Article

A Study on Human Extinction Caused by Technology and the Path to Self-Salvation

Hu Jiaqi*
*Correspondence: Hu Jiaqi, hujiaqi1122@163.com

Abstract

Science and technology are a double-edged sword. They can benefit humanity, yet they can also destroy it. The stronger their capacity to benefit humanity, the greater their power to destroy it. Since the Industrial Revolution in the mid-eighteenth century, science and technology have developed exponentially, yet humanity's rational attitude toward science and technology has remained in a primitive state. I term this phenomenon evolutionary imbalance. Today, high-risk technologies have pushed humanity to the dangerous brink of self-extinction.

Since entering university in 1979, the author have been engaged in systematic research on technological risks for nearly half a century. Several major conclusions and a series of minor ones have emerged from this research. This article elaborates on only two of the important conclusions: first, "the continued development of science and technology will certainly soon lead to the extinction of humanity—at most within two or three hundred years, and at the earliest within this century"; second, "to save humanity from extinction, it is necessary to realize the Great Unification of humanity and strictly control science and technology."

In 2007, my 800,000-character systematic research monograph Saving Humanity was published. In that book, my research conclusion was that the emergence of extinction-level means would not take more than 200 years, and that the eruption of extinction forces would not exceed 300 years thereafter. At the same time, I repeatedly emphasized that this was an extremely conservative estimate and that the reality is far more severe. Therefore, my conclusion was that "if science and technology continue to develop as they do today, they will certainly lead to the extinction of humanity—at most within four or five hundred years, and at the earliest within one hundred years," and I made it clear that "this is a definitive conclusion, not merely a possibility." (Hu, 2007a)

As early as January 2007, when Saving Humanity was still in proof, I sent it, together with The Open Letter to 26 Leaders of Mankind, to 25 leaders of major powers as well as the Secretary-General of the United

Nations, in order to present this research conclusion and convey the gravity of the issue. Today, less than twenty years later, the development of science and technology has proven even faster than I imagined at the time. In fact, not long afterward, in the second edition of Saving Humanity, I revised the time-frame for the extinction of humanity by science and technology to "at most within two or three hundred years, and at the earliest within this century." (Hu, 2019) Although "this century" also falls within "a hundred years," this phrasing conveys a stronger sense of urgency. So, if this research conclusion is valid, what issue today could be more urgent and more important than solving this one? (Hu, 2025)

Over the years, I have devoted myself to using various methods to make people take this issue seriously, but with little effect. This paper presents the method of argumentation behind this conclusion, as well as the solution and pathway for addressing this urgent problem. The former is presented in two parts. The first part, "The Argument That the Continued Development of Science and Technology Will Certainly Lead to the Extinction of Humanity," adopts the method of "Extinction Path Analysis + Defense Limit Testing." The foundation of this analysis rests on five axioms, which can also be regarded as common sense—since axioms are inherently infallible, and common sense is naturally easy to understand. The reasoning process proceeds step by step through five key dimensions, demonstrating the inevitable connection between the continued development of science and technology and human extinction. The second part, "A Time Assessment of the Extinction of Humanity by Science and Technology," is a comprehensive analysis based on the general laws of technological development along the timeline, combined with the current state of development of today's frontier technologies. The latter is addressed in the third part, "The Solution and Realization Pathway for the Overall Survival Crisis of Humanity."

All of these ideas were already present in Saving Humanity nearly twenty years ago. This paper revisits them today, incorporates new developments in science and technology from recent years, and draws on the research achievements of relevant scholars. Its purpose is to provide theoretical support and a practical warning for humanity to face technological risks squarely and adopt urgent countermeasures. (Hu, 2018)

Keywords

Continued development of science and technology, Extinction of humanity by science and technology, Human extinction, Extinction within one hundred years, Extinction within this century, Frontier technologies, Logic of argumentation, Time assessment, The Great Unification of humanity, Awakening movement, Social reconstruction, Non-competitive society, Peaceful society, Friendly society, Equitably prosperous society

1. Introduction

Since the birth of human civilization, the exploration of the unknown and the innovation of tools have always been the core driving forces of human progress. From stone axes and clubs in the Stone Age, to swords, spears, and halberds in the age of cold weapons; from the

steam engine of the Industrial Revolution, to today's nuclear weapons, AI, gene editing, and nanotechnology, science and technology have irreversibly reshaped humanity's modes of production, ways of life, and even the very logic of human survival. (Stavrianos, 2003) In the public imagination, technological progress is positively correlated with human well-being. Every breakthrough in science and technology is regarded as a sign that human civilization is moving to a higher level: it has alleviated humanity's problems of food and clothing, extended human life expectancy, and expanded humanity's living space. It seems that as long as science and technology continue to develop, humanity will achieve eternal prosperity and progress.

However, this "technology-supremacist" understanding is precisely what conceals the fatal risks hidden behind technological development. Based on nearly half a century of systematic research into technological risks, I state responsibly here: as long as science and technology continue to develop in this way, they will certainly soon lead to the extinction of humanity—at most within two or three hundred years, and at the earliest within this century. This conclusion is not alarmist, nor is it subjective speculation. Rather, it is derived from the five axioms of constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, and demonstrated through the scientific method of "Extinction Path Analysis + Defense Limit Testing," (Hu, 2008) and has received strong recognition and resonance from some top scientists.

At present, the speed of technological development has far exceeded humanity's capacity for control. The failure to prevent the proliferation of nuclear weapons, the rapid iteration of AI technology, the disorderly breakthroughs in gene editing, and the potential risks of nanotechnology are all constantly approaching the critical point of human survival. Nearly twenty years ago, in the first edition of Saving Humanity, I initially proposed an extinction-time assessment of "at most within four or five hundred years, and at the earliest within one hundred years." (Hu, 2007a) With the accelerated development of science and technology, this assessment was revised only a few years later to "at most within two or three hundred years, and at the earliest within this century." (Hu, 2019) This revision was not a panic-driven guess, but based on the actual speed of technological development, the potential threats of frontier technologies, and humanity's indifference and inaction toward technological risks.

This paper aims to present, in a comprehensive and systematic manner: (1) the core argument and time assessment concerning the proposition that "the continued development of science and technology will certainly soon lead to the extinction of humanity"; and (2) an in-depth analysis and rational design of "the solution and realization pathway for the overall

survival crisis of humanity." The full text is divided into three parts. The first part uses the method of "Extinction Path Analysis + Defense Limit Testing", grounded in the five axioms— constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, and proceeds through five key steps to demonstrate the inevitable connection between the continued development of science and technology and human extinction. The second part, based on the general laws of technological development along the timeline, combines the current state of frontier technologies such as AI, genetic engineering, nanotechnology, and nuclear weapons, while drawing on the research achievements of some well-known scholars, to verify the core conclusion that extinction will occur "at most within two or three hundred years, and at the earliest within this century." (Hu, 2007c) At the same time, it warns humanity that the unrestrained development of science and technology has become the greatest cause of human extinction, and that facing risks squarely and restricting the disorderly development of technology is the only way for humanity to achieve holistic survival. The third part directly addresses this core question concerning the survival or extinction of humanity. Combining an in-depth analysis of humanity's survival predicament, it explores a feasible solution that can fundamentally defuse the crisis and safeguard humanity's holistic survival, thereby identifying the core answer to the problem of humanity's holistic survival and the pathway for realizing it.

The significance of this paper lies in breaking humanity's blind worship of science and technology, awakening humanity's awareness of the survival crisis, promoting a renewed examination of the relationship between technological development and human survival, and proposing solutions and methods for resolving the technological crisis. It provides theoretical reference and a self-rescue pathway for humanity to respond to the risk of technological extinction and to formulate reasonable mechanisms for constraining technological development.

2. Part I: The Argument That the Continued Development of Science and Technology Will Certainly Lead to the Extinction of Humanity

2.1. Method of argumentation: combining extinction path analysis with defense limit testing

Before arguing the core proposition that "the continued development of science and technology will certainly lead to the extinction of humanity," it is necessary first to clarify the rationality and scientific validity of the research method, as well as the authenticity and reasonableness of the underlying assumptions. This proposition faces an inherent research

dilemma: the question of whether science and technology can exterminate humanity cannot be falsified, observed, or repeatedly tested through traditional scientific methods. Moreover, human civilization is the only sample, and the Earth is humanity's only living space. Traditional scientific research methods are therefore difficult to apply to this unique proposition concerning the holistic survival or extinction of humanity. At present, academia generally adopts probabilistic argumentation to assess the risk that science and technology may exterminate humanity. (Bostrom, 2013; Bostrom, 2002; Cotton-Barratt & Ord, 2015; Metaculus, 2018; Ord, et al., 2010; Ord, 2021; Snyder-Beattie, et al., 2019; Tegmark & Bostrom, 2005) However, after nearly half a century of systematic research, I find that this method is entirely inconsistent with the core logic of human survival issues and cannot provide reliable support for the argument. Its lack of scientific validity is reflected in several core aspects, which is also the key reason I have consistently rejected this method.

At present, the mainstream academic approach to probabilistic assessment of technological extinction risk adopts a combined model of "structured expert judgment + Bayesian updating + multi-path scenarios," including structured expert elicitation, Bayesian probability networks, frequency-based physical models, as well as variants such as war-gaming and the doomsday argument. Yet this type of probability calculation has fundamental defects. It may appear to produce precise numerical conclusions, but in reality, it cannot yield a definite answer. First, the sample size is extremely scarce; it lacks a basic statistical foundation. The scientific validity of probability calculation depends on statistical analysis of a large number of samples; the richer the samples, the closer the result is to objective reality. Yet the extinction of humanity by science and technology is a unique proposition: there is only one human civilization on Earth, and no other advanced biological civilization has yet been discovered. There is no reference sample that can be used to calculate the probability of technology exterminating an advanced species. The limitation of a single sample deprives probabilistic argumentation of its scientific foundation from the outset; any probability calculation is merely subjective speculation. Second, the black swan effect is significant, and the risk contains "unknown unknowns." The real extinction risk may come from fields beyond humanity's current cognition and therefore cannot be incorporated into existing assessment models, leaving a fatal gap in the coverage of probability calculation. Third, the values assigned are subjective and broad, and the conclusions lack uniformity. Even top scholars and institutions provide widely divergent probability assessments, and even the assessment of the same scholar may change over time. For example, Toby Ord of the University of Oxford believes that the probability of technology causing human extinction within 100 years is 1/6, or 16.7% (Ord, 2021); Elon Musk believes that the probability of AI

causing human extinction is about 20% (PowerfulJRE, 2025); a survey of more than 2,700 AI researchers shows that the median probability of AI causing human extinction within a century is 5% to 10% (AI Impacts, 2023); and Max Tegmark believes that the probability of AI losing control exceeds 90% (Tegmark, 2018; Tegmark & Bostrom, 2005); some AI safety researchers argue that once superintelligence emerges, goal misalignment is almost inevitable—i.e., P≈100% (Yudkowsky & Soares, 2025; Yudkowsky, 2022). The most representative case is Geoffrey Hinton, a godfather figure in the field of AI. His assessment of AI losing control has changed several times. He once suggested that AI could lose control within "5 to 10 years," and also stated that the probability of losing control within 30 years is 10% to 20% (BBC Radio 4, 2024; Milmo, 2025). Such changes are not logical contradictions, but repeated emphases on the urgency of technological loss of control. Yet this precisely demonstrates the unreliability of probabilistic argumentation: even top scholars cannot provide stable and unified probabilities, let alone the fact that such probabilities are built on a single sample and subjective value assignments. (Ord, et al., 2010)

More importantly, probabilistic argumentation ignores the uniqueness and irreplaceability of human survival. The value of the holistic survival of humanity is infinite. Whether the probability of technology exterminating humanity is one in ten thousand or one in a million, once multiplied by the infinite value of survival, the final risk cost approaches infinity (Bostrom, 2003). This is something humanity absolutely cannot bear. This method not only fails to reflect the seriousness of technological extinction risk accurately, but also fails to provide reasonable guidance for humanity's survival decisions. Instead, it gives vested interests in technology a reckless excuse: states develop high-risk technologies in pursuit of power, enterprises develop high-risk technologies for high profits, and the public relies on high-tech products for convenience. Everyone remains immersed in an atmosphere of apparent peace and prosperity, while ignoring the precarious technological crisis behind it.

Although the proposition that science and technology may exterminate humanity cannot be falsified, observed, or repeatedly tested through traditional scientific methods, the extinction risk itself can be partially verified and assessed through reasonable means. On this basis, after nearly half a century of accumulated research, I have repeatedly refined a qualitative research method combining "Extinction Path Analysis + Defense Limit Test." Grounded in five axioms, or common sense (since axioms are inherently infallible, and common sense is naturally easy to understand)—constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, the method reaches its conclusion through a five-step deduction. Although this method appears to be qualitative analysis, it in

fact arrives—through clear logical deduction and real-world verification—at a quantitative and certain conclusion. It is far more rigorous and persuasive than probabilistic argumentation.

"Extinction Path Analysis" refers to sorting out the complete logical chain by which technological development leads to human extinction. It clarifies how technological development step by step enhances destructive capability, approaches the extinction threshold, and ultimately triggers an extinction catastrophe. It clearly presents the specific pathway from technological development to human extinction, making an originally abstract extinction risk perceptible and analyzable. This pathway analysis is not subjective speculation, but an objective logic derived step by step from the historical facts of human technological development, its current state, and its inherent laws.

"Defense Limit Testing" by contrast, begins with humanity's defensive capability against technological extinction risk. It assumes that humanity can construct defense systems and develop countermeasures to the greatest possible extent, and then tests the limits of that defensive capability. Even if humanity can develop countermeasures against various extinction-level means, the development of such countermeasures still requires time. During this time gap, humanity will be exposed to extinction risk. At the same time, as extinction-level means continue to increase in number and destructive power, humanity's defensive capability will ultimately fails to keep pace with the development of technological destructive capability. The defensive limit will ultimately be broken, and an extinction catastrophe will ultimately be unavoidable.

In sum, Extinction Path Analysis clarifies the core logic of "how technological development leads to human extinction," while Defense Limit Testing verifies the inevitable conclusion that "humanity cannot prevent such extinction." Combined, the two proceed step by step with rigorous logic. They conform to the laws of historical development as well as the actual state of today's technological development, thereby providing solid argumentative support for the core proposition that "the continued development of science and technology will certainly lead to the extinction of humanity."

2.2. Steps of argumentation: a five-step deduction showing how the continued development of science and technology leads to human extinction

Based on the research method of "Extinction Path Analysis + Defense Limit Testing," and grounded in the five axioms of constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, the author divides the entire argument into five key steps, seeking to proceed step by step and gradually demonstrate the inevitable

connection between the continued development of science and technology and human extinction.

2.2.1. Step one: technological development inevitably leads to the enhanced destructive power of tools

One of the core essences of technological development is the continuous innovation and upgrading of human tools, and the innovation and upgrading of tools inevitably lead to the continuous enhancement of their destructive power. This is a matter of historical common sense that requires no complex proof, and one of the basic laws of technological development. From the origin of human civilization to the present day, every major advance in science and technology has been accompanied by a leap in the destructive power of tools. This enhancement has shown a clear progressive trend and has never been reversed.

In the primitive Stone Age, human tools consisted mainly of natural objects such as stones and clubs. These tools were mainly used to chase wild animals and obtain food, and they were also used in conflicts among humans. At that time, however, the destructive power of tools was extremely limited: the lethality of stones depended on human physical strength, while the attack range and power of clubs were also very limited. A conflict could usually cause only individual casualties and could not threaten the survival of human groups. At that stage, the level of technological development was extremely low, the destructive power of tools was minimal, and the risks to human survival came mainly from the natural environment rather than from technological tools created by humans themselves.

With the preliminary development of science and technology, humanity entered the age of cold weapons. The materials of tools were upgraded from natural stones and clubs to metal products, and cold weapons such as knives, swords, spears, and arrows became the main instruments of attack. Compared with Stone Age tools, the destructive power of cold weapons increased significantly: the hardness and sharpness of metal far surpassed those of stone; knives and swords could easily cut through the human body inflicting fatal injuries; and bows and arrows made long-distance attacks possible, expanding the range of attack. Yet the destructive power of tools at this time still had obvious limitations. Each cold weapon could target only a single opponent. Although the casualties of a war were far greater than in the Stone Age, they remained within a controllable range. Such weapons could not achieve large-scale group destruction, let alone threaten the survival of humanity as a whole.

The invention of gunpowder was an important turning point in the development of human science and technology, marking the entry of tool destructiveness into a new level. The emergence of gunpowder gave rise to explosives, artillery, and other firearms. It changed the way humans fought wars and thoroughly enhanced the destructive power of tools: the

explosion of a single bomb could kill or injure multiple targets within a given area, while the range and power of artillery far exceeded those of cold weapons and could inflict devastating strikes on large crowds and buildings. At this stage, tools already possessed large-scale killing capability. The casualties of a war could reach tens of thousands or even hundreds of thousands, and the destructive power of technological tools had begun to pose a certain threat to the survival of human groups.

At the beginning of the twentieth century, Einstein proposed the mass-energy equation, laying the theoretical foundation for the development of nuclear weapons. The destructive power of human science and technology achieved a leap. The emergence of nuclear weapons completely changed the survival landscape of humanity. The explosion of a single nuclear weapon can destroy a large city and kill or injure hundreds of thousands of people. The global stockpile of nuclear weapons is sufficient to destroy Earth many times over and blast human civilization back to the Stone Age. At this point, technological tools already possessed the ability to destroy humanity on a large scale. The destructive power of technology had been upgraded from “individual killing” and “group killing” to “large-scale destruction.”

It should be noted here that nuclear weapons still cannot exterminate humanity, because nuclear weapons release energy at a single point. Even if all the nuclear weapons in the world were detonated, humanity’s strong survival capacity would mean that a small number of people could still escape the disaster, and humanity would still have the opportunity to start over. Moreover, when a means must be used thousands or tens of thousands of times in order to exterminate humanity, it is no longer an extinction-level means. An extinction-level means refers to a means that leaves not a single person alive, including those who use such means and their families, all of whom would cease to exist in such destruction. To offer a simple analogy: there are more than eight billion people in the world, but the number of bullets and shells stored globally far exceeds eight billion. If distributed evenly among every person, they could also exterminate humanity, but this situation would not occur. Therefore, ordinary guns and artillery are of course not extinction-level means. So-called extinction-level means refer only to those means that can exterminate humanity by being used once or twice, or only a few times.

Thus, from the Stone Age to the nuclear age, the history of technological development is a history of the constantly increasing destructive power of tools. This increase is not accidental, but an inevitable trend of technological development. The core of science and technology is to enhance humanity’s ability to transform nature and the world. This ability includes both the capacity to benefit humanity and the capacity to destroy humanity. The enhancement of the capacity to benefit humanity is inevitably accompanied by the

enhancement of the capacity to destroy humanity; the two complement each other and cannot be separated. Therefore, technological development inevitably leads to the enhanced destructive power of tools. This is an objective law already verified by history and still continuing today, and it is also the basic premise for arguing that "the continued development of science and technology will certainly lead to the extinction of humanity."

2.2.2. Step two: since the industrial revolution, technological development has rapidly approached extinction-level capability

The trend that technological development leads to enhanced destructive power of tools progressed relatively slowly before the Industrial Revolution—indeed, it can be said to have developed extremely slowly. Before the Industrial Revolution, human technological development mainly relied on the accumulation of experience and lacked systematic scientific theoretical guidance. (Dampier, 2020) The speed of technological progress was very slow, and the increase in the destructive power of tools also displayed a gradual pattern. From the Stone Age to the age of cold weapons, and then to the initial development of firearms, tens of thousands of years passed. The increase in destructive power remained limited and never approached the capability to exterminate humanity.

The onset of the Industrial Revolution completely changed the speed and rhythm of technological development. Starting in the mid-eighteenth century, technological development entered an accelerated phase and began to show exponential growth. In just over two hundred years, humanity developed science and technology to an unprecedented level, and the destructive power of tools also increased rapidly, swiftly approaching the capability to exterminate humanity. This accelerated development stemmed mainly from two core factors: first, the deep integration of scientific theory and technological practice; and second, humanity's unlimited pursuit of and investment in technological progress.

Before the Industrial Revolution, scientific theory and technological practice were separated from each other. The development of scientific theory was mainly aimed at exploring the laws of nature, while the development of technological practice was mainly aimed at solving practical problems in production and daily life. The two did not form an effective linkage. As a result, technological progress was extremely slow, and the increase in the destructive power of tools was greatly restricted. After the Industrial Revolution, humanity gradually recognized the guiding role of scientific theory in technological practice and began to apply it to technological research and development, thereby achieving the deep integration of science and technology. Each breakthrough in scientific theory could be rapidly transformed into technological achievements, promoting the innovation of tools and the enhancement of destructive power. At the same time, problems encountered in technological

practice could in turn promote the further development of scientific theory, forming a cycle of "theoretical breakthrough—technological transformation—practical feedback—theoretical refinement," thereby driving science and technology to develop at an exponential rate.

The exponential growth of technological development enabled the destructive power of tools to achieve leaps within a short period of time. As mentioned above, before the Industrial Revolution, humanity spent tens of thousands of years raising the destructive power of tools from stone implements to firearms. After the Industrial Revolution, however, humanity needed only a little more than two hundred years to move from firearms to nuclear weapons, achieving a qualitative leap in destructive power. The emergence of nuclear weapons marked the point at which humanity's technological destructive power had already approached the capability to exterminate humanity.

Since the beginning of the twenty-first century, the speed of technological development has further accelerated. The rapid iteration of frontier technologies such as AI, genetic engineering, and nanotechnology has further enhanced the destructive power of tools, bringing them ever closer to the capability to exterminate humanity. Today, many scientists are deeply concerned about the further development of these frontier technologies. This concern is no longer merely about "how many people can be injured or killed," but about "whether these technologies can exterminate humanity."

Take genetic engineering as an example. The rapid development of gene-editing technology enables humanity to modify the code of life at will. It can be used to treat hereditary diseases and improve human health, but it can also be used to create superviruses and genetic weapons. Once a supervirus is released, and if it can spread rapidly around the world with no effective treatment, it may cause mass human death and even exterminate humanity. The development of nanotechnology enables humanity to create nanorobots. These nanorobots can enter the human body and destroy human cells, and they can also be used for large-scale destruction. Once they lose control, they will pose a fatal threat to human survival. If nanorobot technology achieves further breakthroughs and enables unlimited replication, it could not only exterminate humanity but also destroy the entire Earth. The development of AI, especially the emergence of artificial general intelligence, may allow machines to break free from human control, form autonomous consciousness, and then launch attacks against humanity. With their powerful computing and execution capabilities, the possibility of exterminating humanity would be very high.

What is even more critical is that the speed of technological development is still accelerating. With continual breakthroughs in scientific theory, the constant upgrading of technological research and development methods, and increasing human investment in

science and technology, the speed of technological development will only become faster, not slower. This acceleration means that the destructive power of tools will increase also at an ever-faster rate, moving humanity ever closer to the extinction threshold. Therefore, since the Industrial Revolution, the acceleration of technological development has caused humanity's technological destructive power to rapidly approach extinction-level capability. This is a key step in arguing that "the continued development of science and technology will certainly lead to the extinction of humanity."

2.2.3. Step three: as long as science and technology develop out of control, extinction-level means will inevitably emerge

Based on the arguments of the previous two steps, a clear conclusion can be drawn: technological development inevitably leads to the enhanced destructive power of tools, and since the Industrial Revolution, the accelerated development of science and technology has caused this destructive power to rapidly approach extinction-level capability. The third step further argues that as long as science and technology continue to develop out of control, the destructive power of tools will certainly continue to increase until extinction-level means emerge. This is an inevitable process in which quantitative change leads to qualitative change, and it is also determined by the inherent laws of technological development.

"Extinction-level means" refer to technological means capable of completely exterminating humanity. Once such means emerge, their use can lead to the total extinction of humanity. At present, the nuclear weapons, genetic engineering, nanotechnology, and AI technologies that concern us already possess large-scale killing capability and are only one step away from the ability to exterminate humanity. However, they have not yet reached the level of "extinction-level means."

But this does not mean that these technologies will never reach the extinction level. As long as science and technology continue to develop out of control, the accumulation of quantitative change will ultimately trigger qualitative change, and the emergence of extinction-level means will be inevitable. This is because technological development is a process of constantly pushing limits. Humanity's pursuit of science and technology is boundless. As long as there are no effective constraints, humanity will continue to drive technological breakthroughs and ceaselessly enhance the destructive power of tools. Even if today's nuclear weapons, genetic engineering, nanotechnology, and AI technologies cannot exterminate humanity, with the continued development of science and technology, more powerful and more destructive technological means will certainly emerge. These more powerful means will sooner or later acquire the capability to exterminate humanity. This is a simple principle of quantitative change leading to qualitative change.

What deserves even greater vigilance is that once the earliest extinction-level means emerge, the process will not stop there. More and stronger extinction-level means will continue to appear. With the continued development of science and technology, the types of extinction-level means will become increasingly numerous, their power will become increasingly great, and they will show an obvious trend: they will increasingly tend to become obtainable by certain enterprises, or even by individual scientists.

At present, most high-tech means, especially technological means with large-scale killing capability, are mainly in the hands of states. States possess strong financial resources, talent, and technological capabilities, enabling them to develop and control these high-tech means. Relatively speaking, state behavior can still be constrained through international conventions, diplomatic negotiations, and other methods, rendering the risks relatively controllable. However, as science and technology continue to develop, the threshold for technological research and development will continue to fall, and R&D costs will continue to decline. Some powerful enterprises, and even some top individual scientists, may be able to develop extinction-level means merely with the help of a computer or in their own laboratories.

Compared with nations, the behavior of enterprises and individuals is far more difficult to constrain. The core goal of enterprises is to pursue profit; in pursuit of profit, they may ignore the risks of R&D, or even deliberately develop destructive technological means. The behavior of individuals is highly subjective and uncertain. Once some psychologically extreme individuals who seek revenge against society obtain extinction-level means, they may actively use them, leading to the extinction of humanity. Individual human behavior is very difficult to control. In every era, there have always been people who dare to commit extreme evil. When science and technology develop to the point where individuals can obtain extinction-level means, the survival risk faced by humanity will rise sharply. (Bostrom, 2019)

Therefore, as long as science and technology continue to develop out of control, the destructive power of tools will continue to increase, and the emergence of extinction-level means will be inevitable. Moreover, such means will become increasingly numerous and accessible. This is the core link in arguing that “the continued development of science and technology will certainly lead to the extinction of humanity”: the emergence of extinction-level means is the precondition for human extinction. Once this precondition is established, the extinction of humanity becomes only a matter of time.

2.2.4. Step four: extinction-level means cannot be offset, and the defensive limit will ultimately be broken

Even if extinction-level means have already emerged, this does not necessarily mean that

humanity will certainly be exterminated. If humanity can develop effective countermeasures capable of offsetting the destructive power of such means, it might be possible to avoid an extinction catastrophe. Therefore, the fourth step focuses mainly on humanity's defensive capability against extinction-level means. Through "Defense Limit Testing," it verifies both whether humanity's defensive capability can keep pace with the development of technological destructive capability, and whether it can offset the threat posed by extinction-level means.

Through analysis, it becomes clear that extinction-level means have one core characteristic: they cannot be offset. Or, even if they can be offset, an irreparable time gap remains, and that time gap is sufficient to cause the extinction of humanity. Specifically, this can be divided into two situations.

The first situation is that some extinction-level means cannot be offset at all. Once the destructive power of certain technological means is released, it cannot be offset; humanity can only passively bear the consequences. Take nuclear weapons as an example. Once a nuclear weapon explodes, nuclear energy is released instantaneously, causing large-scale casualties, environmental pollution, and ecological destruction. Such destruction is irreversible. Humanity cannot use any means to offset the harm caused by a nuclear explosion. The dead cannot be brought back to life; the polluted environment cannot be restored in a short period of time; and the damaged ecosystem cannot be quickly repaired. Another example is the loss of control of nanorobots. Once nanorobots begin to replicate in large numbers and destroy human cells or the natural environment, humanity currently has no effective means to stop them quickly and can only passively endure the destructive consequences they bring. Such completely unoffsettable extinction-level means, once used, would inflict a fatal blow on humanity and could even lead to human extinction.

The second situation is that some extinction-level means can be offset, but there exists an irreparable time gap. Some extinction-level means may, in theory, have their threats offset through the development of countermeasures. However, the development of countermeasures requires time and a process. During this time gap, humanity will be exposed to extinction risk. If extinction-level means are used during this period, humanity will be unable to respond in time and will ultimately face extinction.

Take superviruses manufactured through genetic engineering as an example. In theory, humanity could develop vaccines and therapies to offset the threat of superviruses. However, the development of vaccines and drugs requires multiple stages, including clinical trials and mass production, and usually takes months or even years. By contrast, a supervirus may spread extremely quickly. Once released, it may spread across the world within weeks or even days, causing mass infection and death. Before vaccines and drugs are successfully developed,

humanity would have no effective means of response and could only passively become infected and die. By the time vaccines and drugs are developed, the vast majority of humanity may already have died, or humanity may already have been exterminated.

What is even more critical is that, with the continued development of science and technology, the types of extinction-level means will become increasingly numerous. The accumulation of such means will sharply increase humanity's defensive pressure. Humanity's future is extremely long, and every extinction-level means has an "unoffsettable window period." Before countermeasures are successfully developed, humanity will be exposed to the threat of that extinction-level means. As the types of extinction-level means increase, these window periods will overlap with one another, forming an infinitely enlarged "total window period." During this total window period, humanity may face an extinction catastrophe at any time.

Even if we assume that, in the future, humanity can develop countermeasures against all extinction-level means and offset all threats posed by them, the development of such countermeasures will still require time, and window periods will always exist. Humanity's future is extremely long. Over this long span of time, as long as extinction-level means are used once during a window period, humanity will face the fate of extinction. The longer time goes on, the higher the probability that such an "accidental event" will occur. Ultimately, humanity's defensive limit will be broken, and an extinction catastrophe will be unavoidable.

This is the core logic of "Defense Limit Testing": even if humanity can develop countermeasures to the greatest possible extent and offset the threat posed by extinction-level means, it cannot avoid the existence of window periods. The long duration of humanity's future infinitely enlarges such window periods. As long as science and technology continue to develop, and as long as extinction-level means continue to emerge, humanity's defensive capability will never be able to keep pace with the development of technological destructive capability. The defensive limit will ultimately be broken, and an extinction catastrophe will ultimately occur. This is a key support for the argument that "the continued development of science and technology will certainly lead to the extinction of humanity."

2.2.5. Step five: the eruption of extinction forces is inevitable

Through the previous four steps of argumentation, it has become clear that the continued development of science and technology will inevitably lead to the emergence of extinction-level means, while humanity's defensive capability cannot offset the threat posed by such means, and the defensive limit will ultimately be broken. The fifth step of the argument further argues that once extinction-level means emerge, the eruption of extinction forces is unavoidable. Such an eruption may originate from malicious human use, laboratory

negligence, or the careless use of technological products. Whichever pathway occurs, it may lead to the extinction of humanity.

The eruption of extinction forces stems from two core factors: first, the uncertainty of science and technology; and second, the evil side of human nature. These two factors interact with each other and jointly determine the inevitability of the eruption.

First, science and technology are uncertain. Human scientific research is, in essence, an exploration of unknown fields. We cannot fully predict all the consequences of technological development, nor can we fully control all the properties of technological products. Often, technological products we consider "good" may ultimately bring "bad" consequences; technological means that we consider "safe" may ultimately become tools that destroy humanity. This uncertainty means that technological development always carries potential risks. When technology reaches the extinction level, the risks brought about by such uncertainty may turn into extinction catastrophes.

For example, when Freon was first invented, it was regarded as a safe and efficient refrigerant and was widely used in household appliances such as refrigerators and air conditioners. At that time, humanity did not realize its dangers. However, as Freon became widely used, people gradually discovered that it damages Earth's ozone layer, creates holes in the ozone layer, allows large amounts of ultraviolet radiation to reach Earth, increases the risk of skin cancer in humans, and damages ecosystems. Although Freon did not reach the level of extinction-level means, its example fully illustrates the uncertainty of science and technology. We cannot predict all the consequences of technological development. When technology reaches the extinction level, the risks brought by this uncertainty may directly lead to the human extinction.

Second, the evil side of human nature determines that at any time there will be people who seek to commit extreme evil. Human nature is complex. It contains both good and evil; it contains both the desire to pursue beauty and happiness and the desires of greed, jealousy, revenge, and destruction. In every era, there are people who are psychologically extreme, who seek revenge against society, or who pursue destruction. Such people will not retreat in the face of powerful technological means; rather, they may become even more fanatical because of them. When they obtain extinction-level means, they may actively use them, destroying humanity in order to satisfy their extreme desires.

Based on the two core factors above, the eruption of extinction forces mainly has three pathways. These three pathways are independent of one another, yet also interconnected. If any one of them occurs, it may lead to human extinction.

The first pathway is the malicious use of extreme means by extremely evil individuals.

This is the most direct pathway to the eruption of extinction forces. After extinction-level means emerge, some psychologically extreme individuals who seek revenge against society will try every possible way to obtain such means and then use them actively, destroying humanity to vent their dissatisfaction and retaliate. This situation is the most difficult for humanity to guard against. Individual behavior is highly concealed and sudden. We cannot predict who will use extinction-level means, nor can we prevent them in advance from obtaining such means. Once this occurs, humanity will be unable to respond in time and will ultimately face extinction.

The second pathway is laboratory negligence, which leads to the eruption of extinction forces. Extinction-level means are all high-end technologies, and high-end technologies are even more uncertain. Human scientific research is always an exploration of unknown fields. Since it is an exploration of unknown fields, there are many things that are not yet known, and accidents may occur. For example, in a gene-editing laboratory, scientists may accidentally cultivate a supervirus and then unintentionally release it, causing the supervirus to spread worldwide. In a nanotechnology laboratory, scientists may accidentally cause nanorobots to lose control, thereby inflicting a destructive blow on humanity. In an AI laboratory, scientists may accidentally allow AI to develop autonomous consciousness, after which it may launch attacks against humanity. Such laboratory negligence may be accidental, but as science and technology continue to develop and as the types of extinction-level means become increasingly numerous, the probability of such accidental events will become increasingly high. At any time, such events may ignite extinction forces.

The third pathway is the careless use of technological products, which leads to the eruption of extinction forces. Many technological products, when first invented, are regarded as safe and beneficial. Yet in the course of long-term use, they may bring destructive consequences. For example, the careless use of Freon led to the destruction of the ozone layer; the careless use of AI may lead to machines losing control; and the careless use of gene-editing technology may lead to genetic contamination. When technology reaches the extinction level, the consequences of careless use will be even more severe and may even lead to human extinction. Such careless use may result from human negligence, or from humanity's insufficient understanding of the dangers of technological products. Whatever the cause, it may lead to the eruption of extinction forces.

In sum, after extinction-level means emerge, whether through the malicious use by extremely evil individuals, laboratory negligence, or the careless use of technological products, the eruption of extinction forces may occur, and such eruption is unavoidable. This is the final link in arguing that "the continued development of science and technology will

certainly lead to the extinction of humanity": once extinction forces erupt, humanity will be unable to avoid the fate of extinction.

2.3. Summary of the argument: the inevitable connection between the continued development of science and technology and human extinction

Through the method of "Extinction Path Analysis + Defense Limit Testing," grounded in the five axioms of constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, and the step-by-step argumentation of the five steps above, a clear and rigorous conclusion can be reached: the continued development of science and technology will certainly soon exterminate humanity. This is a definitive conclusion, not merely a possibility.

In my view, the logical chain of the entire argument is clear and rigorous: technological development inevitably leads to the enhanced destructive power of tools (Step One); since the Industrial Revolution, the accelerated development of science and technology has caused this destructive power to rapidly approach extinction-level capability (Step Two); as long as science and technology continue to develop out of control, the emergence of extinction-level means is inevitable (Step Three); extinction-level means cannot be offset, and humanity's defensive limit will ultimately be broken (Step Four); ultimately, the eruption of extinction forces is unavoidable, and humanity will eventually be exterminated (Step Five). These five steps are closely interconnected and proceed progressively. They conform to the laws of historical development, correspond to the actual state of today's technological development, and are grounded in the essence of human survival and the characteristics of human nature. They therefore possess strong scientific validity and persuasive power.

It must be emphasized that this argument does not deny the value of science and technology, nor does it advocate the complete abandonment of science and technology. (Hu, 2007c) The wide global dissemination of existing safe and mature scientific technologies would be fully sufficient to ensure that humanity has ample food and clothing. Humanity must not be too greedy. The core purpose of the argument above is to awaken humanity's awareness of the survival crisis and to make humanity realize that the unrestrained development of science and technology has already become the greatest cause of human extinction. Only by stopping the unrestrained development of science and technology and imposing effective constraints on science and technology can humanity avoid the fate of extinction.

3. Part II: A Time Assessment of the Extinction of Humanity by Science and

Technology

3.1. Logic of the time assessment: taking theoretical breakthroughs as the main thread and combining them with the current state of high-risk technologies

Having established that "the continued development of science and technology will certainly lead to the extinction of humanity," an even more urgent core question follows: if science and technology continue to develop without restraint, how long will it take for them cause human extinction? The answer to this question is crucial for humanity's understanding of the urgency of survival risk, and serves as an important temporal reference for formulating countermeasures. If extinction lies far in the future, humanity still has room for buffer; if extinction is imminent, humanity must take emergency action immediately, or it will be too late to reverse the situation.

Since I began systematically researching technological risks in 1979, nearly half a century of deep engagement has continually refined my understanding of this issue. In 2007, my monograph Saving Humanity was published. At that time, the conservative assessment I proposed was as follows: the emergence of extinction-level means would not take more than 200 years, and the eruption of extinction forces would not take more than another 300 years thereafter. (Hu, 2007a) As early as January 2007, when the book was still in sample-copy form, I sent it together with The Open Letter to 26 Leaders of Mankind to 25 leaders of major powers and the Secretary-General of the United Nations. At that time, the timeframe I stated was that "the extinction of humanity by science and technology would occur at most within four or five hundred years, and at the earliest within one hundred years." (Hu, 2025) I believed this should have been sufficient to make global decision-makers aware of the seriousness of the issue.

However, in less than twenty years, the speed of technological development has far exceeded my expectations, and the originally conservative time assessment can no longer match the current pace of technological iteration. Therefore, in the second edition of Saving Humanity, I revised the time assessment for technological extinction to "at most within two or three hundred years, and at the earliest within this century." (Hu, 2019) Although "this century" also falls within "a hundred years," this phrasing conveys a stronger sense of urgency. This revision was not subjective speculation, but an objective conclusion reached after rigorous deduction, based on the current state of technological development, the risks of frontier technologies, humanity's indifference toward those risks, and my long-term research into the laws of science and technology and the cycles of theoretical breakthroughs. It has also been corroborated by the views of some top scholars and entrepreneurs.

The core logic of this time assessment is to take breakthroughs in scientific theory as the main thread and combine this with the current state of today's high-risk technologies.

Only when science and technology develop to a certain height will means capable of exterminating humanity emerge. Yet the height of science and technology always depends on breakthroughs in scientific theory. Scientific theory is the source of technological development; without theoretical breakthroughs, there can be no leap in technology, still less any major increase in destructive power. Therefore, the time assessment must take breakthroughs in scientific theory as its main thread. This is where the core logic lies.

Looking across the history of human science and technology, especially since the Industrial Revolution, every major breakthrough in scientific theory has promoted the rapid development of technology and then increased the destructive power of tools. (Wu, 2022) When theory develops to a certain level, the technological products guided by it will possess the ability to exterminate humanity. This is self-evident. Therefore, the speed of theoretical breakthroughs determines when extinction-level means will emerge, and in turn determine the time of human extinction.

Before the Industrial Revolution, human technological development relied on accumulated experience and lacked systematic theoretical guidance. Progress was slow, and the increase in destructive power was gradual. (Dampier, 2020) From the Stone Age to the age of cold weapons, and then to the initial development of firearms, tens of thousands of years passed, yet destructive power never approached the extinction level. The core reason was the absence of systematic scientific theory as support; technology could not achieve a qualitative breakthrough.

The advent of the Industrial Revolution brought about the deep integration of scientific theory and technological practice, forming a cycle of "theoretical breakthrough—technological transformation—practical feedback—theoretical refinement," (Wu, 2022) and driving science and technology to develop exponentially. From that point onward, breakthroughs in scientific theory became the core driving force of technological development, as well as the key to enhancing the destructive power of tools.

It should be made clear that theoretical breakthroughs in different disciplines have different degrees of impact on technology. As the core of the natural sciences, physics has the most significant effect on technological development and the enhancement of destructive power. From Newtonian mechanics promoting the development of firearms, to relativity and quantum mechanics giving rise to nuclear weapons, and then to quantum mechanics supporting breakthroughs in AI and nanotechnology, every revolution in physical theory has brought a leap in destructive power. Therefore, in this time assessment, breakthroughs in

physical theory serve as the core reference point, while the laws of other disciplines are also considered in order to infer the emergence time of extinction-level means.

The time assessment is not fixed is dynamically adjusted according to the speed of technological development. The faster science and technology develop, the shorter the cycle of theoretical breakthroughs becomes, the earlier extinction-level means emerge, and the earlier human extinction may occur. In 2007, I proposed "at most within four or five hundred years, and at the earliest within one hundred years," and later revised this to "at most within two or three hundred years, and at the earliest within this century." (Hu, 2019) The core reason was that the speed of technological development had far exceeded expectations, and the cycles of theoretical breakthroughs and technological iteration had both surpassed my judgment at the time.

In addition, breakthroughs in scientific theory are showing a trend of cross-disciplinary integration. Today's four major frontier technologies—AI, genetic engineering, nanotechnology, and nuclear weapons—all rely on cross-disciplinary theoretical support. Nuclear weapons rely on nuclear physics, relativity, and related theories; genetic engineering is built on molecular genetics and related foundations; nanotechnology depends on quantum mechanics and related theories; and AI integrates mathematical logic, neuroscience, and other disciplines. This cross-disciplinary integration accelerates the speed of theoretical breakthroughs, shortens the time required for technological transformation, and further brings forward the emergence of extinction-level means.

3.2. The laws and cycles of breakthroughs in physical theory: the core reference for the emergence time of extinction-level means

Modern physics has undergone three revolutionary theoretical breakthroughs, each of which has promoted a leap in technology and a major increase in destructive power. By analyzing the cycles of these three breakthroughs and combining them with the law of accelerated technological fission, it is possible to reasonably infer the timing of the fourth theoretical breakthrough, and thereby infer the emergence time of extinction-level means. (Toynbee, 2001) This is one of the core bases of the time assessment.

3.2.1. The three revolutionary theoretical breakthroughs of modern physics and their cycles

The three revolutionary theoretical breakthroughs of modern physics laid the foundation for technological development at different stages and promoted a stepwise increase in destructive power. Their breakthrough cycles reveal clear patterns and provide an important reference for inferring the timing of the fourth breakthrough.

The first revolutionary theoretical breakthrough, or Level One, was Galilean physics, which emerged from the late sixteenth century to the early seventeenth century. It marked the beginning of modern physics, broke the constraints of theology, and established the scientific research method of "observation, experiment, and reasoning." However, because the Industrial Revolution had not yet arrived and humanity had not yet recognized the significance of applying theory, this theoretical achievement remained largely at the academic level for a long time. Its substantive impact on technological practice was delayed by nearly a century. At that time, the destructive power of tools was still dominated by cold weapons and could not threaten the survival of human groups.

The second revolutionary theoretical breakthrough, or Level Two, was Newtonian mechanics, established in the late seventeenth century. Its core consisted of the law of universal gravitation and the three laws of motion. It completely changed humanity's understanding of the universe and provided theoretical support for the deepening Industrial Revolution. The transformation of this theory into technology took about 100 years. With the widespread application of machinery such as the steam engine, Newtonian mechanics was applied to manufacturing, railway construction, and other fields. At the same time, it promoted the rapid improvement of firearm technology: the range and power of artillery and rifles increased substantially, and destructive power was upgraded from individual killing to large-scale killing, beginning to threaten the survival of human groups.

The third revolutionary theoretical breakthrough, or Level Three, was relativity and quantum mechanics, which emerged in the early twentieth century and were founded by scientists such as Einstein and Bohr. (Tegmark, 2017) They laid the theoretical foundation for high-end technologies such as nuclear weapons and aerospace technology. Unlike the previous two breakthroughs, the transformation of these theories into technology took only about 40 years. By this time, humanity had already attached importance to the application of theory and had invested large amounts of funds and talent to promote the rapid implementation of theoretical achievements.

The mass-energy equation of relativity, $E=mc^2$, revealed the conversion relationship between mass and energy and provided the core theory for the development of nuclear weapons. Quantum mechanics revealed the laws governing the motion of microscopic particles; its theories of atomic nuclear structure, chain reactions, and related fields supported the development of atomic bombs and hydrogen bombs. In 1945, the United States dropped atomic bombs on Japan, marking the transformation of Level Three theoretical achievements into actual weapons. Human destructive power achieved its second great leap and had already approached the extinction level. The global stockpile of nuclear weapons is sufficient to

destroy Earth many times over. Although it cannot completely exterminate humanity—because humanity has a strong capacity for survival and some people could escape the disaster—it is already only one step away from extinction.

To summarize the cycles of the three breakthroughs: the interval from the first to the second breakthrough was less than 100 years, while the interval from the second to the third breakthrough was about 200 years. Although the breakthrough cycle appears to have lengthened. (Toynbee, 2001) In reality, however, higher-level theoretical breakthroughs require the accumulation of more intermediate achievements. This reflects the increasing depth and difficulty of theoretical development, rather than a slowdown in the speed of breakthroughs. The rate at which scientific and technological achievements accumulate has always been accelerating.

3.2.2. Time inference for the level four revolutionary theoretical breakthrough in modern physics

The Level Four revolutionary theoretical breakthrough will be the key to the emergence of extinction-level means. Technologies guided by Level Three theories, such as nuclear weapons, have already approached the extinction level. As a higher-level breakthrough, Level Four theory will break through the boundaries of current physical understanding, and technological products guided by it will inevitably possess the ability to exterminate humanity with ease. Therefore, inferring the timing of the Level Four theoretical breakthrough is central to inferring the emergence of extinction-level means and the time of human extinction.

By combining the cycles of the first three breakthroughs with the law of accelerated technological fission, the timing of the Level Four theoretical breakthrough can be reasonably deduced. Level Three theories—relativity and quantum mechanics—were established in the early twentieth century, about 120 years ago. The cycles of the previous two breakthroughs were less than 100 years and about 200 years, respectively. Meanwhile, the law of accelerated technological fission—disciplinary subdivision and accelerated accumulation of achievements—will compensate for the shortage of intermediate achievements needed for higher-level theoretical breakthroughs and shorten the breakthrough cycle.

After Level Three theories were established, science and technology entered a stage of explosive development. Physics, biology, and other disciplines increasingly integrated; applications such as quantum computing and aerospace technology continued to advance. The new problems and discoveries arising from these practices, in turn, have promoted theoretical refinement and accumulated a large number of intermediate achievements for the Level Four theoretical breakthrough.

By conservative estimation, the breakthrough cycle of Level Four theory will be about

150 to 200 years, meaning that a breakthrough may occur in the middle or late twenty-first century. The time required to transform theory into technology is expected to shorten to ten or twenty years—Level Three theory took about 40 years, and the speed of technological transformation continues to accelerate. Therefore, extinction-level means guided by the Level Four revolutionary theory of physics may emerge in the middle or late twenty-first century, or in the early twenty-second century.

However, this inference is based on a conservative assumption regarding the emergence pattern of extinction-level means guided by Level Four revolutionary physical theory. The current speed of technological development has far exceeded expectations. The iteration speed of frontier technologies such as AI and genetic engineering, as well as the speed of theoretical breakthroughs in physics, biology, and other disciplines, are all accelerating significantly. For example, the application of quantum mechanics has extended to quantum computing, gene-editing technology has broken through the limits of traditional treatment, and AI is rapidly advancing toward strong AI and artificial general intelligence. Yet all of these high-risk technologies are based only on Level Three revolutionary physical theory.

It must be emphasized that theoretical breakthroughs in other disciplines, such as biology and computer science, need not wait for Level Four physical theory. They may independently promote the emergence of extinction-level means. For example, breakthroughs in molecular genetics may promote the development of genetic weapons, and breakthroughs in neuroscience may accelerate the realization of artificial general intelligence. All of these will further shorten the emergence time of extinction-level means.

In just the past decade or so, humanity's awareness of technological extinction risk has shifted from indifference to concern. Behind this shift lie the rapid development of science and technology and the continual manifestation of risks. This change in awareness also confirms my judgment: the speed of technological development has far exceeded expectations, and the emergence of extinction-level means is very likely to occur earlier than the conservative estimate.

3.2.3. Inference of the eruption time after the emergence of extinction-level means

After extinction-level means emerge, humanity will not necessarily be exterminated immediately; the eruption of extinction forces will still require some time. By combining the historical patterns of the use of extreme means with the current characteristics of science and technology, it is possible to reasonably infer the eruption time. This is also an important part of the time assessment.

In human history, every kind of large-scale killing means has often been used shortly after it was developed. Nuclear bombs were put into war immediately after they were

successfully developed in 1945; genetically modified biological toxins, after more than thirty years of development, have become key objects of prevention by various countries. Yet extinction-level means are fundamentally different from ordinary extreme means: the user would also perish together with everyone else. Therefore, rational actors would not actively use them, and states would strengthen control over them, thereby prolonging the eruption time.

Even so, the eruption of extinction forces remains unavoidable, and the longer time passes, the higher the probability of eruption becomes. There are three core reasons.

First, the research and development threshold for extinction-level means will continue to fall, making them easier to obtain. At present, high-risk technologies are mainly controlled by states and are relatively controllable. However, as science and technology develop, the threshold and cost of research and development will continue to decline. Powerful enterprises, and even top individual scientists, may be able to develop extinction-level means merely through laboratories or computers. Enterprises pursue profit, and individual behavior—such as that of extremists—is difficult to constrain. Once they obtain extinction-level means, the risk of eruption will rise sharply.

Second, the uncertainty of science and technology may lead to accidental eruption. Humanity's exploration of science and technology is marked by the unknown, and its consequences cannot be fully predicted. After extinction-level means emerge, factors such as laboratory operational errors, program failures, and inaccurate understanding of the properties of technological products may cause leakage or loss of control. The eruption time cannot be predicted.

Third, the evil side of human nature determines that eruption is inevitable. Human nature is complex and contains extreme desires such as greed and revenge. Once extremists obtain extinction-level means, they may actively use them to satisfy their own extreme demands.

Overall, after extinction-level means emerge, the eruption time will not be very long. Means controlled by states may erupt within several decades; means controlled by enterprises or individuals may erupt within a few years or even less; and accidental eruption may occur at any time. Therefore, if extinction-level means emerge in the middle or late part of this century, their eruption will very likely occur within this century. Even if they emerge in the early twenty-second century, their eruption may also occur in the first half of the twenty-second century.

3.3. The current development status of today's frontier technologies: the core basis for revising the time assessment

The assessment that human extinction may occur "at the longest in two or three hundred

years, and at the shortest within this century" is based not only on the patterns of theoretical breakthroughs in physics, but also on the current development status of four major frontier technologies: AI, genetic engineering, nanotechnology, and nuclear weapons. Their rapid iteration and uncontrolled breakthroughs are accelerating the emergence of means capable of causing extinction, shortening the timeline toward human extinction. Their speed of development has far exceeded expectations and is constantly approaching the critical threshold of human survival.

These four major technologies rely on their own core theories and integrate achievements from foundational disciplines. They have moved from hypothesis toward real-world application, demonstrating the driving power of theory while also exposing the enormous risks of technological loss of control. The following sections discuss their current development status, underlying theories, and extinction risks, explaining why the time assessment is increasingly leaning toward "at the shortest within this century."

3.3.1. The rapid iteration of ai technology: accelerating the emergence of extinction-causing means

AI technology is currently the fastest-developing frontier technology. From narrow AI, such as voice assistants and autonomous driving, to strong AI with human-level cognition, and then to artificial general intelligence that surpasses human cognition, its pace of iteration has far exceeded expectations. It is not only transforming human production and daily life, but also bringing enormous extinction risks and accelerating the emergence of extinction-causing means. All of this is inseparable from the support of its underlying theories.

AI relies on multidisciplinary theories such as mathematical logic, probability theory, linear algebra, and neuroscience. It is an interdisciplinary product of mathematics, computer science, and life sciences. (Russell & Norvig, 2023; Russell, 2015) Mathematical logic and Turing's theory of computability define the boundaries of machine intelligence; mathematical theories support neural network training and data modeling; information theory and cybernetics build the framework for system operation; and neuroscience provides inspiration for deep learning, driving breakthroughs in large models.

At present, narrow AI has been widely applied. Although it lacks autonomous consciousness, it has already demonstrated powerful capabilities: autonomous driving carries risks of loss of control, and generative AI can be used to create false information and develop lethal tools. Even more worrying is that the development of strong AI and artificial general intelligence has entered an accelerated stage.

Geoffrey Hinton, a godfather-level figure in the AI field, has warned that the probability of generative AI losing control within 30 years is 10% to 20% (BBC Radio 4, 2024; Milmo,

2025), and he has also suggested that loss of control may occur within 5 to 10 years. These changing figures reflect both the uncertainty and rapidity of AI development, while also highlighting the severity of the risk of loss of control. Based on the current pace of iteration, a breakthrough in artificial general intelligence by the end of this century is by no means impossible.

Artificial general intelligence would possess cognitive and executive capabilities surpassing those of humans. Once it loses control, it could rapidly master all of humanity's scientific and technological knowledge, develop extinction-causing means unimaginable to humans, or even directly launch extinction attacks. Its speed of learning and evolution would far exceed that of humans, leaving humanity with little ability to fight back. (Bostrom, 2012; Tegmark, 2018; Yudkowsky & Soares, 2025; Yudkowsky, 2022)

Even more alarming is the continuous lowering of the threshold for AI research and development. The spread of open-source platforms allows more enterprises and individuals to participate in AI development. Extremists may use AI to develop destructive tools, such as viral code or drone swarms, further increasing the risk of loss of control. (Bostrom, 2011)

In addition, the combination of AI with other frontier technologies accelerates the emergence of extinction-causing means. AI plus genetic engineering could rapidly design super viruses; AI plus nanotechnology could precisely control nanorobots; AI plus nuclear weapons could enable autonomous launch. These cross-disciplinary applications greatly accelerate the development speed and destructive power of extinction-causing means.

In summary, AI is highly likely to achieve a breakthrough in artificial general intelligence by the end of this century. Once it loses control, extinction-causing means may emerge within a short period of time, leading to human extinction. This is an important basis for the judgment that extinction could occur "at the shortest within this century."

3.3.2. The uncontrolled breakthroughs of genetic engineering: the accelerated emergence of potential extinction risks

Genetic engineering alters the code of life by editing genes. It can be used to treat genetic diseases and improve health, but it can also be used to create super viruses and genetic weapons, bringing enormous extinction risks. At present, genetic engineering has entered a stage of uncontrolled breakthroughs. Countries are competing in research and development while neglecting risks, causing potential extinction risks to emerge at an accelerated pace. Its development relies on the underlying theories of life sciences.

Genetic engineering is built upon theories such as molecular genetics, the DNA double-helix structure, and the central dogma. The theory that DNA is the core genetic material provides the fundamental basis for gene editing; the theory of the DNA double-helix

structure makes artificial manipulation of genes possible; the central dogma clarifies the chain of genetic function, supporting the development of transgenic technologies and genetic medicines; and the theory of the universality of the genetic code breaks down species barriers, ensuring the feasibility of cross-species genetic recombination.

The improvement of these underlying theories has driven the rapid development of gene-editing technology. The popularization of CRISPR-Cas9 has made gene editing precise, efficient, and low-cost, greatly lowering the threshold. Small laboratories and even individuals can now conduct related research. This technology can cut DNA like “scissors,” enabling the insertion, deletion, and replacement of genes. It can be used for medicine, but it can also be used to develop genetic weapons and super viruses.

Although the international community has imposed certain constraints on gene editing, such as prohibitions on human embryo editing and the development of genetic weapons, these measures have failed to prevent uncontrolled breakthroughs. Some scientists have crossed ethical boundaries to conduct embryo editing, while some countries have secretly developed genetic weapons, ignoring extinction risks.

Genetic weapons can selectively attack specific populations and specific vital organs, or create super viruses with high transmissibility, high pathogenicity, and no effective treatment. Such viruses could spread through air, water, and other channels. Their concealment, selectivity, and destructiveness would far exceed those of nuclear weapons. Once released, they could rapidly cause mass death and even human extinction.

For example, scientists could use gene editing to modify smallpox or Ebola viruses, enhancing their transmissibility and pathogenicity. They could also develop weapons targeting specific genes, achieving “precise extinction.” Once such means are maliciously used or accidentally leaked, they would trigger a global pandemic and lead to human extinction.

In addition, careless use of gene editing could cause genetic contamination. Edited genes might accidentally enter the natural environment, disrupting ecological balance and affecting human survival. For example, crop genes could hybridize with wild plants and cause species extinction, or viral genes could leak and trigger epidemics.

At present, genetic engineering is developing extremely rapidly. The popularization of CRISPR-Cas9 has sharply reduced the difficulty of developing super viruses and genetic weapons. At this pace, such means capable of extinguishing humanity may appear by the end of this century. Once used, human extinction could very likely occur within this century.

3.3.3. The potential threat of nanotechnology: an important carrier of extinction-causing means

Nanotechnology studies the microscopic properties and controllable processing of matter

at the nanoscale, namely 10^{-9} meters. It is applied in fields such as medicine and the military. However, its "microscopic nature" and "replicability" bring enormous extinction risks, making it an important carrier of extinction-causing means. Its development relies on underlying theories from multiple disciplines.

Nanotechnology is based on theories such as quantum mechanics, mesoscopic physics, and surface effects. Quantum mechanics governs the properties of matter at the nanoscale, such as quantum confinement effects. Surface effects give nanomaterials extremely strong catalytic properties. Small-size effects explain abrupt changes in material properties, and theories of intermolecular forces support the precise construction of nanomolecules.

At present, nanotechnology has achieved certain breakthroughs. Nanomaterials are used for drug delivery and tumor treatment, and progress has also been made in the development of nanorobots, which may precisely manipulate human cells. However, behind these technologies lie fatal risks: the microscopic nature and replicability of nanorobots make them ideal carriers of extinction.

Nanorobots are only nanometer-sized. They could easily penetrate the skin, enter the bloodstream, and even infiltrate cells, precisely destroying human tissues. At the same time, they could self-replicate. Once they lose control, they would reproduce wildly like viruses, destroying the human body and the natural environment, ultimately leading to human extinction. Their destructive power would grow exponentially, far exceeding that of nuclear weapons.

To give a hypothetical example of nanotechnology application: nanorobots could be used to clean blood vessels, kill cancer cells, or transform matter. To improve efficiency, they might be designed to self-replicate: one nanorobot replicates into ten, ten into a hundred, and within a short period of time they could reach hundreds of millions. But once this replication loses control, it would cause an uncontrollable catastrophe.

If nanorobots replicated without limit, they could use the atoms of human bodies and the Earth as raw materials, rapidly consuming both humans and the planet, like billions of locusts destroying everything. Some scientists believe their replication could be controlled through programming, but this idea is overly idealistic. A program malfunction, human negligence, or extremists deliberately omitting control mechanisms—any one of these scenarios could lead to human extinction.

At present, the threshold for nanotechnology research and development continues to decline. Countries are focusing on its application value while neglecting risks, and effective constraints are lacking. Based on the current speed of development, uncontrolled nanorobots may appear by the end of this century. Once leaked or maliciously used, they would deliver a

devastating blow to humanity.

3.3.4. The upgrading and iteration of nuclear weapons: an existing potential extinction risk

Nuclear weapons, as a core technology guided by Level-Three theory, already possess tremendous destructive power. Although they cannot yet completely extinguish humanity, they are undergoing upgrading and iteration, bringing existing potential risks and pushing humanity closer to the critical threshold of extinction. Their development relies on core theories such as nuclear physics and relativity. Their pace of upgrading reflects the urgency of technological development and provides additional basis for the time assessment.

The underlying theories of nuclear weapons include the mass-energy equation of relativity, $E = mc^2$, the theory of atomic nuclear structure, and the theory of chain reactions. The mass-energy equation established the energy logic of nuclear weapons; chain reaction theory supported the development of the atomic bomb; and nuclear fusion theory supported the development of the hydrogen bomb, enabling a leap in the destructive power of nuclear weapons.

At present, nuclear weapons have entered a stage of upgrading and iteration. Nuclear powers are developing hypersonic nuclear strike capabilities, which can break through missile defense systems, as well as miniaturized nuclear weapons, which lower the threshold for use. These developments greatly increase the risk of nuclear use.

Even more alarming is the intensified risk of nuclear weapons proliferation. The spread of nuclear technology has enabled more countries to acquire development capabilities. Some countries secretly conduct research and development despite opposition, creating the possibility that nuclear weapons may fall into the hands of extremist organizations or individuals, greatly increasing the probability of use. Although the large-scale use of nuclear weapons may not completely extinguish humanity, it would cause mass death, environmental pollution, and ecological collapse.

In addition, the combination of AI and nuclear weapons further increases the risk. AI could be used for autonomous nuclear launch and precision strikes. Once AI loses control, it may autonomously activate nuclear weapons. Missile defense systems are not 100% reliable; once they fail, nuclear strikes cannot be intercepted, and the consequences would be devastating.

In summary, the upgrading, iteration, and proliferation risks of nuclear weapons are pushing humanity closer to the critical threshold of extinction. Their combination with other frontier technologies would form even more destructive extinction-causing means, further shortening the timeline toward human extinction.

3.4. Time assessment: at most within two or three hundred years, and at the earliest within this century

Integrating the above arguments—from beginning with a qualitative research method to arriving at a quantitative and definite conclusion; from the core logic of breakthroughs in physical theory, to the deduction of the laws of theoretical breakthroughs, and then to the current state of the four major frontier technologies—after nearly half a century of deep research and repeated deduction, the following conclusion can be reached: the continued unrestrained development of science and technology will inevitably lead to the extinction of humanity, and the time of this extinction will be at most within two or three hundred years, and at the earliest within this century. This conclusion is built upon scientific laws, the current state of technology, and the essence of human nature. It is an objective warning about the risks to human survival.

"At most within two or three hundred years" is a conservative deduction. Its core basis lies in the breakthrough cycle of Level Four physical theory and the time required for technological transformation. Level Three theory was established in the early twentieth century, about 120 years ago. Referring to the cycles of the first three breakthroughs and the law of accelerated technological development, the breakthrough cycle of Level Four theory will be about 150 to 200 years, meaning that it may occur in the middle or late twenty-first century. Adding another ten or twenty years or so for technological transformation, extinction-level means may appear in the middle or late twenty-first century, and at the latest no later than the early twenty-second century. Even under the most ideal state of control, considering state competition, the lowering of technological thresholds, and the complexity of human nature, the eruption of extinction forces would still require several decades. Therefore, by conservative estimation, the longest timeframe will not exceed two or three hundred years.

Two or three hundred years is not distant. Compared with the tens of thousands of years of human civilization, it is extremely short; compared with the hundreds of millions of years that nature should have given humanity, it is only an instant. Dinosaurs survived on Earth for 160 million years, and nautiluses have survived for 530 million years, while humanity's ability to cope with nature is even stronger than theirs. Compared with the lifespan of an individual, two or three hundred years is merely three or four generations. More importantly, this assessment is based on the conservative assumption that "the speed of science and technology remains unchanged and humanity adopts certain constraints." In reality, however, science and technology are accelerating, while humanity remains indifferent to risks. This "maximum timeframe" is therefore very likely to be shortened.

"At the earliest within this century" is an urgent judgment based on the current state of

frontier technologies and the risk of loss of control. It is also the core reason for my revision of the assessment. The disorderly breakthroughs of the four major frontier technologies have accelerated the emergence of extinction-level means. Their speed has far exceeded the expectations of 2007 and far exceeded humanity's capacity for control.

In terms of AI technology, weak AI has already become widespread, while the research and development of strong AI and artificial general intelligence are accelerating. Many well-known scholars have warned of the risk of AI losing control. It is not impossible for artificial general intelligence to be achieved by the end of this century. Once it loses control, extinction-level means may rapidly emerge.

In terms of genetic engineering, CRISPR-Cas9 technology has lowered the threshold for gene editing, and the difficulty of developing superviruses and genetic weapons has decreased. By the end of this century, such extinction-level means may emerge. Once they are used or leaked, human extinction may occur within a short period of time.

In terms of nanotechnology, progress has been made in the development of self-replicating nanorobots. With the lowering of research and development thresholds and the absence of effective constraints, if out-of-control nanorobots emerge by the end of this century, their devouring capacity will rapidly destroy humanity and Earth.Together with the upgrading and proliferation of nuclear weapons, as well as the cross-integration of the four major technologies—such as AI plus nuclear weapons and AI plus genetic engineering—the eruption of extinction forces will be further accelerated, greatly increasing the possibility that humanity will face an extinction crisis within this century.

It must be emphasized that this assessment is not intended to create panic, but to face a real crisis. Over nearly half a century, I have revised my conclusion from "at most within four or five hundred years, and at the earliest within one hundred years" to the current conclusion: "at most within two or three hundred years, and at the earliest within this century." This is not the result of a research deviation, but because technological development has far exceeded expectations and the risks have become increasingly urgent. Hawking once said that human civilization will either be destroyed in disasters or continue through interstellar exploration. But in my view, if science and technology continue to develop without restraint, humanity will not have the opportunity for interstellar exploration. The urgency of the situation no longer allows humanity to delay.

The core significance of the time assessment is that it provides a temporal reference for countermeasures. If we recognize the risk of "at the earliest within this century," then we must act immediately: stop the unrestrained development of science and technology, establish a unified global risk-control mechanism, and curb the research, development, and application of

high-risk technologies. If we ignore the risk, human extinction will no longer be a distant prophecy, but an imminent crisis.

4. Part III: The Solution and Realization Pathway for the Overall Survival Crisis of Humanity

Truth is often something people do not like very much, yet truth cannot be avoided.

The preceding text has used the scientific method of "Extinction Path Analysis + Defense Limit Testing," grounded in the five axioms of constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, and combined with the laws of breakthroughs in physical theory and the current development of today's four major frontier technologies, to clearly demonstrate the core conclusion: the unrestrained and continued development of science and technology will inevitably soon exterminate humanity, at most within two or three hundred years, and at the earliest within this century. This crisis is neither a distant science-fiction imagination nor a panic-driven subjective claim. Rather, it is an inevitable conclusion reached after nearly half a century of systematic research and repeated deduction, based on the inherent laws of technological development, the complex essence of human nature, and the real predicament of global governance.

Since it has already been made clear that "the unrestrained development of science and technology will certainly lead to human extinction," the logic for resolving the crisis becomes clear and firm: if development will inevitably and soon lead to extinction, then avoiding extinction requires restriction. That is, we must restrict the disorderly research, development, and proliferation of high-risk technologies, abandon the blind pursuit of technological benefits, and strictly place technological development within the core framework of "serving humanity's holistic survival." This logic appears simple and straightforward, yet it is extremely difficult for all humanity to accept. The core problem is that there is a fundamental contradiction between the current development logic of human society and the needs of the holistic survival of humanity.

At the practical level, states compete for national strength with science and technology at the core, treating technological capability as a key bargaining chip for maintaining international status, gaining a competitive advantage, and even winning wars. The speed of technological development is directly tied to the rise and fall of states, causing countries to fall into the prisoner's dilemma of "if we do not develop technology, we will fall behind; if we fall behind, we will be beaten." No state is willing to take the initiative to restrict its own technological development. (Armstrong, et al., 2013) At the enterprise level, technological

research and development are treated as core competitiveness. Companies rely on high-tech products and technological breakthroughs to obtain high profits and expand market share, binding technological progress deeply to corporate survival. Their pursuit of short-term technological benefits far exceeds their consideration of humanity's long-term survival risks. At the level of ordinary people, the public has long grown accustomed to the convenient life brought by technological products. From daily travel and communication to food, clothing, housing, and transportation, technology has become deeply integrated into every aspect of life. People's dependence on technological convenience makes it difficult for them to accept the proposal to "restrict technological development."

Yet it must be made clear that the holistic survival of humanity overrides all. This is the premise and bottom line of all development logic. Without the existence of humanity itself, the national strength of states, the profits of enterprises, and the convenience enjoyed by the public would all lose any meaning, and all technological benefits would be reduced to nothing. In the face of the risk of human extinction, any short-term or local interest is insignificant. To face this core proposition concerning humanity's survival or extinction, and to explore a feasible path that can fundamentally defuse the crisis and safeguard humanity's perpetual development, has become the most urgent and important task of the present. We must not continue to indulge in the short-term benefits brought by technology while ignoring the imminent survival crisis.

Based on nearly half a century of in-depth analysis of humanity's survival predicament, systematic research into technological risks, and practical exploration, I believe that the core answer to the problem of humanity's holistic survival lies in realizing the Great Unification of humanity, breaking down the barriers of the current division among states, forming a common consensus among all humanity, and using the power of a world regime to strictly restrict the disorderly development of science and technology. (Hu, 2007b) Only safe and mature technological achievements should be disseminated globally to meet humanity's basic production and living needs. At the same time, the social system must be reconstructed, the current competitive social model must be abandoned, and the ultimate goal of humanity's holistic survival must finally be achieved.

A deeper analysis shows that the technological extinction crisis currently facing humanity is rooted in the predicament of global governance under division and in the lack of constraints on technological development. In the pattern of divided states, countries fall into the "prisoner's dilemma" of technological competition. They treat technological capability as the core bargaining chip for competing for national interests and maintaining international status, forming a disorderly competitive situation in which each side tries to outpace the

others. No state is willing to take the initiative to restrict its own technological development, because otherwise it would face the risk of "falling behind and being beaten." This short-sighted competitive logic will ultimately push all humanity into the abyss of extinction. At the same time, humanity's long-standing blind worship of science and technology and greedy pursuit of its gains have gradually caused technological development to deviate from its essence of "serving human survival." Technology has been reduced to a tool for pursuing interests and competing for advantage, and humanity is moving step by step toward the edge of self-destruction.

Therefore, the key to resolving the crisis lies in breaking down national barriers, forging a human consensus, reversing the disorderly state of technological development, and achieving a fundamental transformation through systematic pathway design: from "technology out of control" to "technology under control," from "state division" to "human unity," and from "survival crisis" to "perpetual development." Specifically, the pathway to resolving humanity's survival crisis and achieving perpetual development can be divided into five interconnected and progressively advancing core steps. These steps are closely linked and organically unified, together forming a complete pathway of "striving in the present—awakening and enlightenment—human unity—technology control—social reconstruction—human perpetuation." This pathway provides practical guidance for humanity to escape the extinction crisis and realize long-term survival.

4.1. Striving in the present: building a temporary line of defense for crisis prevention and control

To resolve humanity's survival crisis, we must both take the long view and plan for the root solution, while also facing the present and acting proactively. As we advance the long-term solution, we must adopt every feasible measure to win more buffer time for humanity. "Striving in the present" is not a stopgap measure, but a necessary part of crisis prevention and control. Its core goal is to slow the research and development of high-risk technologies, reduce the probability of technological disasters, and lay a foundation for the subsequent awakening movement and the unification of humanity.

At present, some scholars around the world have already recognized the risk of technology losing control and have proposed targeted ideas for prevention and control. The "alignment theory" proposed by Yudkowsky and others is a typical example. (Armstrong, 2013; Joy, 2000; Orseau & Armstrong, 2016; Russell, 2014; Russell, et al., 2015; Russell, 2020; Yudkowsky, 2016; Yudkowsky, 2001) This theory seeks to use technical means to keep the goals of AI aligned with humanity's core values, thereby preventing superintelligence from behaving in ways contrary to human interests and thereby reduce the risk of AI losing

control. Objectively speaking, although such theories and practices cannot fundamentally resolve the extinction crisis brought about by the continued development of science and technology—after all, they remain limited to the category of "controlling a single technology," fail to touch the core roots such as the division among states, technological competition, and weaknesses in human nature, and cannot address the comprehensive risks posed by multiple fields of high-risk technologies such as genetic weapons, nanorobots, and synthetic pathogens—it cannot be denied that, in the short term, they can play a certain buffering role. To some extent, they can curb the abuse of some high-risk technologies and buy more time for humanity to advance fundamental solutions. "Having them is better than having nothing." At a time when crisis is imminent, any effort that helps avoid technological disaster deserves attention and implementation.

In addition to supporting such technical explorations of prevention and control, "striving in the present" should also include concrete measures such as promoting the construction of a global consensus on technological ethics, strengthening international regulation of high-risk technological research and development, and restricting investment in certain extremely high-risk technologies. States should temporarily set aside short-term competition for interests and, under the coordination of international organizations such as the United Nations, establish a shared monitoring mechanism for high-risk technological research and development. They should strictly control technological fields that may threaten human survival, resolutely curb malicious research and development and technological abuse, and win precious time for humanity's awakening and unification.

4.2. Awakening and enlightenment: reshaping humanity's cognitive consensus on science and technology

To realize the Great Unification of humanity and technology control, the prerequisite is to promote an awakening movement of all humanity, completely break humanity's blind worship of science and technology, and re-understand the nature and boundaries of technology. For a long time, humanity has regarded science and technology as "the primary productive force," equated technological development with "progress," and one-sidedly pursued the material convenience and economic benefits brought by technology, while ignoring its destructive potential. This cognitive bias is precisely the ideological root of technology losing control. (Yudkowsky, 2008) Therefore, the core task of the awakening movement is to guide all humanity toward a correct view of science and technology, to make clear that "technology is a tool, not an end," and to recognize that technological development must take human survival as its premise, rather than stand above human survival.

This awakening and enlightenment movement is not simple conceptual propaganda, but

a cognitive revolution that spans the globe and reaches deep into the entire population. It must cross the boundaries of state, nation, religion, and culture, and forge a survival consensus among all humanity. First, through education, media, academic exchange, and various other channels, the survival crisis caused by technology losing control should be widely communicated, so that everyone understands that the unrestrained development of science and technology is not "benefiting humanity," but pushing humanity toward the edge of extinction. This will break the entrenched belief that "technological development inevitably brings progress." Second, humanity's value pursuits should be re-examined. Humanity should be guided away from "maximizing material interests" and toward "pursuing holistic survival and universal well-being." (Hu, 2025) Greedy, selfish, and short-sighted thinking should be abandoned, and the core value that "the holistic survival of humanity overrides all" should be established. That is, the value judgment of all human activities should take "the holistic survival and universal well-being of humanity" as its measure. When individual or group interests conflict with the holistic survival of humanity, the former must unconditionally submit to the latter.

In order to promote this awakening and enlightenment movement, I have always practiced what I advocate. Since 1979, I spent 28 years completing the 800,000-character theoretical work Saving Humanity, in which I not only argued this urgent proposition but also proposed a complete solution. Since 2007, I have written twelve open letters successively to leaders of various countries, heads of international organizations, renowned scientists, scholars, and others, sending a total of one million letters and calling on them to take the risk of technological extinction seriously. In 2008 and 2019, I also established two global promotional websites, respectively, using more than ten languages specifically to promote awareness of the technological crisis. After 2010, making use of my certain talent for business, I concentrated all my funds and energy in an attempt to become the world's richest person. In my view, if I could become the world's richest person, I would not only have more funds to promote the cause of humanity, but as the richest person, my words would also carry enough weight to attract people's attention. Although this effort ended in failure and cost me almost all the funds I had at the time, I have no regrets. In 2018, the year before I realized that my plan to become the world's richest person might be difficult to achieve and prepared to stop that project, I founded "Humanitas Ark" — formerly the Save Human Action Organization — and decided to use collective strength to promote the cause of awakening and enlightenment. At present, Humanitas Ark has brought together more than 14 million supporters around the world. I have also continued to spread the ideas of technological safety and the Great Unification of humanity through academic exchanges, public lectures, published articles,

dialogues with renowned scholars, and many other forms.

The various awakening and enlightenment practices I have carried out have made me fully realize that although all of this is extremely difficult, it is not a castle in the air. As long as we continue to exert effort and spread the message widely, we can gradually break down cognitive barriers, make more and more people aware of the common crisis facing humanity, and lay a solid ideological foundation for realizing the Great Unification of humanity.

4.3. The great unification of humanity: the fundamental prerequisite for resolving the crisis

The ultimate purpose of promoting human awakening is to realize the Great Unification of humanity and to use the power of a world regime to control science and technology. This is the fundamental prerequisite for resolving the crisis of technology losing control and ensuring humanity's holistic survival. As stated above, technological competition under the condition of divided states is the core root of the unrestrained development of science and technology. No single country can escape the limitations of self-interest and voluntarily restrict the research and development of high-risk technologies. Existing international organizations such as the United Nations, because they lack unified authority and executive power, are also unable to effectively control global technological development or resolve the "prisoner's dilemma" among states. Therefore, only by realizing the Great Unification of humanity and establishing a world regime that transcends national boundaries can humanity truly stand from the standpoint of all humanity, coordinate the overall situation, control science and technology, and fundamentally defuse the survival crisis.

The Great Unification of Humanity that I advocate is not a simple merger of states, but the establishment of "a peaceful, friendly, equitably prosperous, and non-competitive society." Through a world regime, global resources would be coordinated, human interests would be harmonized, regional barriers and ethnic and religious divisions would be broken down, and the common development of all humanity would be realized. This Great Unification is rooted in the reality of the "global village" brought about by modern transportation and communication technologies, and it is an inevitable choice consistent with the laws of human development. The core function of the world regime would be to abandon national self-interest, take "the holistic survival of humanity" as its highest goal, and at the same time consider "humanity's universal well-being." To this end, it would formulate globally unified rules for technology control, resource allocation plans, and systems of social governance, ensuring that all human activities serve the ultimate goal of humanity's perpetual development.

Realizing the Great Unification of humanity will require a long process. It must cross

many obstacles, including national interests, ethnic cultures, and religious beliefs, and it requires the joint efforts and firm determination of all humanity. In my view, the hardware conditions for the Great Unification of humanity are already in place; "all that is missing is consensus and determination." As long as the awakening movement continues to advance, and as more and more people recognize the highest common interest of humanity, let go of differences, and build consensus, the goal of the Great Unification of humanity will certainly be achieved.

4.4. Technology control: upholding the core bottom line of human survival

After the Great Unification of humanity is realized, the core task will be to strictly restrict the development of science and technology and uphold the core bottom line of human survival. Such restriction does not deny the value of science and technology; rather, it allows technology to return to its essence of "serving human survival," abandons the blind pursuit of high-risk technologies, and realizes the harmonious coexistence of science and technology with human survival. Specifically, the core measure of technology control is to "popularize safe technologies and seal away high-risk technologies."

On the one hand, existing safe and mature science and technology should be widely disseminated throughout the world to guarantee the basic survival needs and quality of life of all humanity. At present, the safe and mature technologies already mastered by humanity are sufficient to solve basic survival problems such as food, medical care, and energy. They can enable all humanity to have ample food and clothing and live and work in peace and contentment. There is no need to pursue additional material interests through the research and development of high-risk technologies. A world regime should coordinate global technological resources, fairly distribute safe and mature technological achievements to all regions, narrow regional development gaps, realize common prosperity for all humanity, and lay a material foundation for humanity's perpetual development and universal well-being.

On the other hand, high-risk technologies and their related theories should be permanently sealed away, even to the point of being forgotten. The high-risk technologies referred to here include not only frontier technologies such as AI, gene editing, nanorobots, and synthetic pathogens, but also the basic scientific theories that can give rise to these high-risk technologies. This is because theory is the root of technological development. Only by sealing away high-risk theories can the research, development, and application of high-risk technologies be curbed at the root. A world regime should establish a strict mechanism for technology control, prohibit any organization or individual from researching, developing, or disseminating high-risk technologies and related theories, centrally seal away and destroy existing high-risk technological achievements, and completely cut off the pathway by which

high-risk technologies threaten human survival. Such sealing away is not "regression," but humanity's rational control over its own destiny. It is for the sake of safeguarding humanity's holistic survival and represents the wisdom of "preventing trouble before it happens."

4.5. Social reconstruction: building the institutional guarantee for perpetual development

Strict technology control requires a corresponding social system as support. Therefore, while realizing the Great Unification of humanity and controlling technological development, it is also necessary to comprehensively reconstruct existing society and establish a new social system adapted to humanity's perpetual development. The core of social reconstruction is to abandon the current social operating model centered on "competition" and establish "a peaceful, friendly, equitably prosperous, and non-competitive society." Such a society will not only help ensure the long-term stability of a unified human society, but also help people obtain a universal sense of well-being. (Bostrom, 2025)

The specific content of social reconstruction is extremely extensive. It is a large theoretical system covering social order, distribution systems, education systems, cultural development, and many other aspects. At the level of social order, a world regime should resolve ethnic and religious divisions, redirect military spending toward people's livelihood, reduce violent conflicts, and build a stable and peaceful social environment. At the level of the distribution system, based on "the Principle of Justice," it should break resource monopolies, regional barriers, and intergenerational selfishness, realize the fair distribution of global resources, narrow the gap between rich and poor, and achieve equitable prosperity for all humanity. At the level of the education system, it should abandon the "technology-supremacist" educational philosophy, generally implement intermediate technical education, take "the holistic survival of humanity," "technological ethics," and "cooperative win-win" as core educational content, and cultivate citizens with a global vision, a sense of crisis, and responsibility. At the level of cultural development, it should promote global cultural integration, cultivate common value identity among humanity, break the narrow belief that "those who are not of my kind must have different hearts," and form a mode of coexistence marked by mutual respect, tolerance, and mutual assistance.

The essence of social reconstruction is the fundamental optimization of the operating model of human society. It is to return social development to its essence of "serving human survival and happiness," provide a solid institutional guarantee for technology control, and ensure that humanity, on the basis of controlling technology and safeguarding survival, can realize universal well-being and holistic survival.

Social reconstruction alone is an enormous subject. I have discussed it relatively

comprehensively in my related works. At the same time, I also candidly acknowledge that, from the perspective of the Principle of Far-sightedness, if things unfold according to nature's arrangement, our species should still have hundreds of millions of years ahead of it. (Hu, 2007b) Many years from now, perhaps even the physical characteristics of our species will change. The plan for social reconstruction will naturally need to be adjusted accordingly as circumstances change.

5. Conclusion

This paper, using the research method of "Extinction Path Analysis + Defense Limit Testing", grounded in the five axioms of constant human nature, the unidirectional escalation of technological destructive power, the irreversibility of group competition, the perpetual lag of defense, and the uncertainty of technology, proceeds step by step through five key stages and fully argues the core proposition that "the continued development of science and technology will certainly soon exterminate humanity." It clearly presents the inevitable connection between the continued development of science and technology and human extinction: technological development inevitably increases the destructive power of tools; after the Industrial Revolution, the accelerated development of science and technology has caused this destructive power to rapidly approach extinction-level capability; development out of control will inevitably give rise to extinction-level means; humanity's defensive limit will ultimately be broken; and the eruption of extinction forces is unavoidable. At the same time, taking breakthroughs in scientific theory as the main thread and combining this with the laws of theoretical breakthroughs in modern physics and the current development of frontier technologies such as AI, genetic engineering, nanotechnology, and nuclear weapons, this paper makes the time assessment that "if science and technology continue to develop without restraint, the time leading to human extinction will be at most within two or three hundred years, and at the earliest within this century." This conclusion is not subjective speculation, but the result of my nearly half a century of deep research into technological risks. It is also corroborated by the views of some top scholars and entrepreneurs, and it is an objective warning reached after rigorous deduction.

In my view, the research in this paper has both important theoretical significance and practical value. At the theoretical level, it breaks the traditional understanding that "technological progress necessarily brings improvement in human well-being," constructs a research system of "Extinction Path Analysis + Defense Limit Testing," abandons the probabilistic argumentation method commonly adopted in academia, and seeks to enrich the theoretical achievements of research on human extinction risk, providing theoretical reference

and methodological inspiration for subsequent related studies. (Hu, 2008)At the practical level, it seeks to awaken humanity's awareness of the survival crisis, warn humanity to face squarely the fatal risks behind technological development, promote humanity's re-examination of the relationship between technological development and human survival, break the blind worship of science and technology, spread the idea that "technology must have boundaries," and provide practical and feasible real-world solutions for humanity to formulate mechanisms for constraining technological development and adopt emergency countermeasures.

Human survival has never been something that can be taken for granted. Today, as science and technology continue to develop without restraint, humanity is standing on the edge of the cliff of self-extinction. The time assessment of "at most within two or three hundred years, and at the earliest within this century" is by no means alarmist. Rather, it is a rational warning about the fate of humanity: the time left for humanity is already running short. (Rees, 2003a; Rees, 2003b; Rees, 2020) Facing technological risks squarely and adopting emergency countermeasures is the only way for humanity to realize holistic survival. Through an in-depth analysis of the laws of technological development, the essence of human nature, and the current state of global governance, this paper makes clear that the roots of the crisis of technology losing control lie in technological competition under the division of states and in humanity's greedy nature. Yet in the face of this ultimate crisis concerning humanity's survival or extinction, humanity is not without a solution.

The rescue pathway I propose—"striving in the present—awakening and enlightenment—human unity—technology control—social reconstruction—human survival"—seeks to explore a feasible way for humanity's holistic survival. This pathway takes the Great Unification of humanity as its fundamental prerequisite, technology control as its core key, social reconstruction as its institutional guarantee, and the awakening of all humanity as its foundational support. It both faces the grave reality of technology losing control and takes into account humanity's survival needs and pursuit of happiness, providing a clear direction of action for the perpetual development of human civilization. In this pathway, "striving in the present" builds a temporary line of defense and wins buffer time for humanity; "awakening and enlightenment" reshapes cognitive consensus and breaks the ideological shackles of technology worship; "human unity" resolves the predicament of division and unites the common strength of all humanity; "technology control" upholds the bottom line of survival and allows technology to return to its essence of serving human survival; and "social reconstruction" optimizes the operating model and builds a new social system adapted to perpetual development.

It must be made clear that realizing this path of rescue will by no means be easy. It requires all humanity to set aside differences of state, nation, and culture, abandon short-term private interests, and form the core consensus that "the holistic survival of humanity overrides all." It requires crossing numerous obstacles and making long-term and steadfast efforts. It requires breaking the entrenched understanding of "technology supremacy," redefining the relationship between technology and humanity, and establishing a rational view of science and technology. We must clearly recognize that this is humanity's only choice: either continue blindly pursuing technological benefits, allow technology to lose control, and ultimately move toward extinction; or awaken proactively, walk forward hand in hand, and safeguard humanity's common future through the Great Unification of humanity and technology control.

Science and technology are a double-edged sword. They can benefit humanity, but they can also destroy humanity. Their ultimate direction depends on humanity's own choice. Since 1979, I have spent nearly half a century deeply researching technological risks, writing the monograph Saving Humanity, sending letters to leaders of various countries, implementing the plan to become the world's richest person, founding Humanitas Ark, and devoting all my efforts to promoting human awakening and the Great Unification of humanity, all for the purpose of awakening every person's sense of crisis. (Hu, 2019; Hu, 2018; ЦзяциХу, 2020) I have always firmly believed that, as members of humanity, each of us has the responsibility to spread the idea of technological safety, participate in the awakening movement, and contribute our own strength to safeguarding humanity's perpetual survival.

The continuation of human civilization has never depended on the efforts of a single country, but on all humanity standing shoulder to shoulder. Only by setting aside differences, building consensus, upholding the core bottom line of "the holistic survival of humanity," strictly controlling technological development, and reconstructing the social system can human civilization escape the crisis of extinction and realize perpetual development in the long river of the universe. May every person clearly recognize the current survival crisis. May all humanity unite in heart and effort, work together, and walk a path of holistic survival that belongs to humanity itself, so that the flame of human civilization may continue forever.

Acknowledgments

Not applicable.

Author contributions

Hu Jiaqi individually undertook all roles including conceptualization, methodology, validation, formal analysis, investigation, resources, data curation, original draft writing, review &

editing, visualization, supervision and project administration for this study, and has read and agreed to the published version of the manuscript.

Funding

Not applicable.

Data availability

Not applicable.

Declarations

Conflict of interest

The author declares that there is no conflict of interest.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

All authors have given their consent.

Additional information

Received: 21 August 2026

Accepted: 24 August 2026

Published online: 3 September 2026

Publisher's Note

References

[1] AI Impacts. (2023). Thousands of AI authors on the future of AI. Journal of Artificial Intelligence Research, 84, Article 9.

[2] Armstrong, S. (2013). Risks and mitigation strategies for Oracle AI. *Studies in Applied Philosophy, Epistemology and Rational Ethics*, 5, 335–347. https://doi.org/10.1007/978-3-642-31674-6_25

[3] Armstrong, S., Bostrom, N., & Shulman, C. (2013). Racing to the precipice: A model of artificial intelligence development. Technical Report.

[4] BBC Radio 4. (2024). Today [Video]. YouTube. https://www.youtube.com/watch?v=gnr0otD4HD0

[5] Bostrom, N. (2002). Existential risks: Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology, 9*(1).

[6] Bostrom, N. (2003). Astronomical waste: The opportunity cost of delayed technological development. *Utilitas, 15*(3), 308–314.

[7] Bostrom, N. (2011). Information hazards: A typology of potential harms from knowledge. *Review of Contemporary Philosophy*, 10, 44–79.

[8] Bostrom, N. (2012). The superintelligent will: Motivation and instrumental rationality in advanced artificial agents. *Minds and Machines, 22*(2).

[9] Bostrom, N. (2013). Existential risk prevention as global priority. *Global Policy*, 4, 15–31.

[10] Bostrom, N. (2019). The vulnerable world hypothesis. *Utilitas*, 31.

[11] Bostrom, N. (2025) Deep Utopia: Life and Meaning in a Solved World. Beijing: CITIC Press.

[12] Cotton-Barratt, O., & Ord, T. (2015). Existential risk and existential hope: Definitions. Future of Humanity Institute Technical Report.

[13] Dampier, W. C. (2020). A brief history of science. China Overseas Chinese Press.

[14] Hu, J. Q. (2007a). Saving Humanity (1st ed., Vol. 1: The extinction of humanity). Tongxin Publishing House.

[15] Hu, J. Q. (2007b). Saving Humanity (1st ed., Vol. 2: The path for humanity). Tongxin Publishing House.

[16] Hu, J. Q. (2007c). The greatest problem. Northeastern University Press.

[17] Hu, J. Q. (2008). On human extinction. Ha Fai Yi Publication Limited.

[18] Hu, J. Q. (2018). Saving Humanity. Canada: FriesenPress.

[19] Hu, J. Q. (2019). Saving Humanity (3rd ed.). South Ocean Publishing House.

[20] Hu, J. Q. (2025). The Twelfth Open Letter to Human Leaders.

[21] Joy, B. (2000). Research priorities for robust and beneficial AI. *Wired, 2000*(4).

[22] Metaculus. (2018). If a global catastrophe occurs, will it reduce the human population by 95% or more? https://www.metaculus.com/questions/9060/#comments

[23] Milmo, D. (2025). AI firms warned to calculate threat of super intelligence or risk it escaping human control. The Guardian.

[24] Ord, T. (2021). The Precipice: Existential risk and the future of humanity. CITIC Press.

[25] Ord, T., Hillerbrand, R., & Sandberg, A. (2010). Probing the improbable: Methodological challenges for risks with low probabilities and high stakes. *Journal of Risk Research, 13*(4), 191–205.

[26] Orseau, L., & Armstrong, S. (2016). Safely interruptible agents. Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence.

[27] PowerfulJRE. (2025). Joe Rogan Experience #2281 - Elon Musk [Video]. YouTube. https://www.youtube.com/live/sSOxPJD-VNo

[28] Rees, M. (2003a). Our Final Century.

[29] Rees, M. (2003b). Our Final Hour: A scientist's warning. Basic Books.

[30] Rees, M. (2020). On the Future: Prospects for humanity. Shanghai Jiao Tong University Press.

[31] Russell, S. (2014). Of myths and moonshine. Edge.org.

[32] Russell, S. (2015). Unifying logic and probability: Recent developments. *Communications of the ACM, 58*(7).

[33] Russell, S. (2020). Human Compatible: Artificial intelligence and the problem of control. CITIC Publishing Group.

[34] Russell, S., & Norvig, P. (2023). Artificial Intelligence: A modern approach. People's Posts and Telecommunications Press.

[35] Russell, S., Dewey, D. & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *AI Magazine, 36*, 105-114. https://doi.org/10.1609/aimag.v36i4.2577

[36] Snyder-Beattie, A. E., Ord, T., & Bonsall, M. B. (2019). An upper bound for the background rate of human extinction. *Scientific Reports*, 9.

[37] Stavrianos, L. S. (2003). A global history: From prehistory to the 21st century. Shanghai Academy of Social Sciences Press.

[38] Tegmark, M. (2017). Our mathematical universe: My quest for the ultimate nature of reality. Zhejiang People's Publishing House.

[39] Tegmark, M. (2018). Life 3.0: Being human in the age of artificial intelligence. Zhejiang Education Publishing House.

[40] Tegmark, M., & Bostrom, N. (2005). How unlikely is a doomsday catastrophe? Nature.

[41] Toynbee, A. (2001). A study of history.

[42] Wu, G. S. The Course of Science (2022). Changsha: Hunan Science and Technology Press.

[43] Ху, Цзяци. Спасение Человечества (2020). Russia: Eksmo.

[44] Yudkowsky, E. (2001). Creating Friendly AI 1.0: The analysis and design of benevolent goal architectures. Machine Intelligence Research Institute.

[45] Yudkowsky, E. (2008). Cognitive biases potentially affecting judgement of global risks. In Global Catastrophic Risks (pp. 91–119).

[46] Yudkowsky, E. (2016). The AI alignment problem: Why it's hard, and where to start. Machine Intelligence Research Institute.

[47] Yudkowsky, E. (2022). AGI Ruin: A list of lethality. https://intelligence.org/2022/06/10/agi-ruin/

[48] Yudkowsky, E., & Soares, S. (2025). If Anyone Builds It, Everyone Dies: Why superhuman AI would kill us all. Little, Brown and Company.