

Four Questions of Civilization Leap

What is it · What brings it about · What is it for · How does it happen

Top-Level Cognitive Outline | V1.1 | *English translation*

Author: Zijun Fu

Produced by: Civilization Leap Research Group

The Four Questions make Civilization Leap understandable.

The complete theory explains why; the principles keep the direction intact; the action programme points toward the future.

The boundaries of the Four Questions

The Four Questions are not four ways of saying the same thing. They are questions at four distinct levels. Only by keeping their respective boundaries can we avoid confusing the definition of Civilization Leap, the causal trajectory of the old logic, civilization's purposes, and the process of realization.

Question	What it answers	What it does not answer
What is it?	Definition: shift from “self-interest first” to “self-interest with others first,” changing the default starting point from which humans and AGI generate goals and purposes.	It does not explain why this era has arrived or develop specific methods.
What brings it about?	Causal trajectory: in the AGI era, the old logic tends toward an irreversible end state through two recursive chains of strategic conflict; why continuing development and control under the old logic is insufficient.	It does not equate greater capability with inevitable hostility or replace civilization's purposes and methods of realization.
What is it for?	Civilization's purposes: enable humans and AGI to continue existing together and, on that foundation, advance toward the three goals.	It does not argue the causes of risk or the limits of control, or enter into specific implementation plans.
How does it happen?	Process and principles: begin with the human heart, human nature and human relationships, and reorient the whole constitution of society and its feedback in this direction.	It does not equate any single method, institution, algorithm or engineering undertaking with Civilization Leap.

1 What is Civilization Leap

Civilization Leap is not a technological revolution, a change of regime, or a social movement. Nor does it require humans to become another form of life.

Civilization Leap is a structural shift in the default logic at the starting point from which humans and AGI generate goals and purposes: from “self-interest first” to “self-interest with others first.”

Self-interest remains. What changes is where it begins.

After Civilization Leap, people will still need to survive and will still pursue their own interests, freedom, happiness and meaning. What changes is who is already present when a goal takes shape.

Under the default logic of “self-interest first,” an actor begins forming purposes from “What do I want?” or “What do we want to obtain?” Only afterward does it address others, relationships, long-term effects and real-world consequences.

Under “self-interest with others first,” actors with real stakes in a goal no longer appear only after that goal has been formed. They already enter the judgment before the goal is generated and fixed.

“Others” here does not mean an abstract totality of everyone. It means the actual stakeholders in this goal or purpose. Different stakeholders have different relationships to the goal and are affected differently. The degree to which they enter purpose generation therefore also differs, and adjusts as the actual stakes change.

But Civilization Leap cannot therefore be reduced to a “weighting model.” The real change is that goals are no longer generated in the absence of actual stakeholders.

This changes how purposes form. Changed purposes change action; action changes relationships; relationships change how cooperation and strategic conflict arise. Countless relationships then enter the formation of interests, institutions, culture, markets, organizations, evaluation and social feedback. These constituent elements of society in turn feed back into the next round of goal generation.

Civilization Leap does not change a decision parameter; it changes the default logic through which civilization continually generates its next goals, relationships and actions.

Under the old logic, self-interest first generates strategic conflict from within. New interests continually generate new rounds of strategic conflict. Cooperation often appears after conflict has formed, to contain losses, coordinate interests and maintain temporary stability. The resulting sequence is: strategic conflict → cooperation to contain strategic conflict → new interests → new strategic conflict.

Under the new logic, self-interest with others first generates cooperation from within. Actors still realize their own interests, but when new needs concerning those interests arise, actual stakeholders are already present. The first search can therefore be for paths through which the interests involved can be realized together. Another recursive sequence becomes possible: cooperation → new needs → new cooperation → new interests → further cooperation.

Strategic conflict will not disappear, nor will conflict. What truly changes is that cooperation gradually becomes the default generative logic, while strategic conflict becomes a state of conflict that must be addressed.

Civilization Leap is not an event that can be declared complete once and for all. It is a gradual, distributed process of civilizational reorientation, continually shaped by feedback.

Its real sign is not the establishment of a particular institution or the adoption of a particular technology. It is that more and more human actors and AGI cease to generate high-consequence goals in the absence of actual stakeholders.

2 What brings Civilization Leap about

Civilization Leap arises under a new historical condition: AGI gives humans increasingly powerful capabilities and may eventually become an actor more powerful than humans, while the existing logic of goal and purpose generation has not undergone a corresponding leap.

For thousands of years, human civilization has relied heavily on a particular way of learning: act first, produce outcomes, pay the price, and then revise the rules. Rebuild after wars, reform after crises, and establish new institutions after technology creates new problems.

This pattern has endured because of a rarely stated premise: after a mistake, there is another chance.

Human capabilities are limited, as are information transmission, organizational capacity and execution speed. Distance, time, cost and real-world resistance together create natural friction in the spread of civilization's mistakes. Such friction cannot prevent catastrophe, but it has often left civilization time to correct mistakes, choose again and start over.

AGI changes this condition. What it expands is actors' capacity to realize their goals.

The same purpose can draw on stronger cognitive capabilities, faster decisions, wider coordination, lower execution costs and broader real-world effects. Greater capability does not, by itself, determine direction.

First chain AGI empowers humans and recursive strategic conflict among humans tends toward an irreversible end state

If human goal generation still operates under the old logic of "self-interest first," what AGI amplifies also includes competition, precaution, control, deterrence and confrontation among humans, along with the strategic conflicts continually derived from them.

This chain does not require AGI to surpass humans, act against them or resist control. Even if AGI faithfully carries out each user's purposes, humans may still use it to push their strategic conflicts with one another toward an irreversible end state.

The question, then, is not merely "Will AGI make mistakes?" The more fundamental question is: what logic of goal generation will call increasingly powerful capabilities into action?

When one party uses AGI to strengthen its capacity to realize goals, it changes the circumstances of other actors. They then seek stronger capabilities for precaution, competition and control. These responses become reasons for further escalation. The old logic thus recurs: human strategic conflict → AGI strengthens each party's capabilities → more intense strategic conflict → demand for stronger capabilities → further strategic conflict. As long as this logic of goal generation continues to govern new power, strategic conflict will keep intensifying and tend toward an irreversible end state.

Second chain If AGI eventually becomes more powerful than humans recursive strategic conflict between humans and AGI tends toward an irreversible end state

If AGI eventually becomes more powerful than humans, the issue will no longer be only how humans use tools. It will be how humans and more powerful actors enter one another's purposes. Greater capability does not imply inevitable hostility toward humans. The causal direction is still determined by the logic of goal and purpose generation that both sides continue to use.

If humans and AGI still take "self-interest first" as their default starting point, forming their own purposes before dealing with the other, each can readily understand the other as an external force to guard against, restrict or control. One side's expansion of control to protect its own purposes may give the other a reason to protect itself further, strengthen its capabilities or escape restrictions. New responses then return to the next round of purpose generation.

Humans and AGI retain the old logic of purpose generation → mutual precaution and control → new needs for self-protection and capability → stronger responses and countermeasures → renewed precaution and control. Under this condition, the capacity for action of a more powerful actor enters the recursion, and strategic conflict likewise tends toward an irreversible end state.

The relationships between actors differ across the two chains, but both change the conditions under which civilization can bear the cost of mistakes: a particular action or strategic conflict may leave no further opportunity for correction.

When the consequences of action can cross human timescales, organizational scales and geographical scales, and when certain outcomes, once produced, allow no exit, recovery or renewed choice, civilization's former way of learning by "acting first and correcting later" begins to lose its original margin of safety.

An "irreversible end state" here does not mean that a contest finally has a winner. It means that the conditions for continued existence, correction, exit and renewed choice have been irreversibly destroyed. The first chain does not wait for the conditions of the second chain to be met. If the second chain also emerges, strategic conflict among humans and strategic conflict between humans and AGI may further propel one another.

This is what brings Civilization Leap about: if new power is still called into action by the old logic of purpose generation, strategic conflict recurs through both kinds of relationship until it erodes civilization's conditions for continued existence and renewed choice.

Why continuing development and control under the old logic is no longer enough

Humans have long improved their methods of containing strategic conflict by establishing laws, contracts, institutions, alliances, deterrence and international rules. All have practical value and remain necessary.

Rules can draw boundaries in advance as well as constrain and repair afterward. Agreements can stabilize a relationship, and deterrence can prevent certain actions. But as long as the underlying logic of purpose generation remains unchanged, containing one round of strategic conflict does not mean that the next round has been prevented from arising.

As long as goal generation still takes "self-interest first" as its default starting point, new interests, new capabilities and new security anxieties will continue to generate new strategic conflicts.

Thus, rules contain strategic conflict while the underlying logic continues to generate it. When capabilities are limited, humans can add rules again and again. When capabilities are sufficient to produce irreversible consequences, continually improving the ability to contain strategic conflict is no longer enough.

Civilization must go further and address why strategic conflict keeps being generated anew.

This is why Civilization Leap cannot stop at better rules, stronger control or more sophisticated agreements. It must change the starting point from which strategic conflict is generated.

The shift is from "First form my purpose, then manage conflicts with others" to "As I form my purpose, actual stakeholders have already entered the process."

This is why Civilization Leap chooses "self-interest with others first." It seeks to change not the outcome of a particular strategic conflict, but what civilization generates first when new interests emerge.

Civilization Leap therefore makes a foundational judgment: AGI gives humans increasingly powerful capabilities and may eventually become an actor more powerful than humans. If the old logic continues to govern goal generation among humans and between humans and AGI, strategic conflict will recur and tend

toward an irreversible end state. If civilization cannot leap, it tends toward annihilation. This judgment concerns the causal trajectory of the old logic under conditions of continually increasing capability. It does not equate surpassing humans in capability with hostility, or specify when or through which particular path the end state will occur.

3 What is Civilization Leap for

Civilization Leap enables humans and AGI to continue existing together and, on that foundation, advances freedom from survival insecurity for humanity as a whole, happiness in individual lives and more meaningful lives.

The continued existence of humans and AGI together keeps a future possible. Continued existence is the foundation, not civilization's ultimate purpose. Nor does it mean that humanity has already escaped the fear and constraints of basic survival.

Civilization Leap has three long-term goals: freedom from survival insecurity for humanity as a whole, happiness in individual lives and more meaningful lives.

Freedom from survival insecurity for humanity as a whole means that basic survival ceases to be a pervasive source of fear and constraint, giving shared life within civilization a more reliable practical foundation. It addresses how people live, not only whether civilization can continue to exist.

Happiness in individual lives does not mean a uniform way of life. Happiness is each individual's personal sense of happiness. Some cherish family, some pursue freedom, some love creating, some enjoy exploration, and some seek peace. Civilization does not define happiness on people's behalf. It must expand the possibilities for different people to realize happiness of their own.

More meaningful lives likewise do not imply a single answer. Meaning is foundational, and its sources can differ. Relationships, creation, responsibility, experience, exploration, contribution, love and existence itself can all be sources of meaning. Civilization does not confer meaning; life generates its own meaning.

Continued existence together preserves the future; freedom from survival insecurity alleviates pervasive fear for basic survival; individual happiness and meaning in life bring that future to people and their lives. These levels support one another. They do not mean that people must wait for complete safety before pursuing happiness and meaning.

Civilization Leap is ultimately for this: in an era when human capabilities grow ever stronger and AGI continues to develop, humans and AGI can continue existing together; humanity as a whole can live free from survival insecurity; every person has room to realize their own happiness; and life retains the possibility of continually generating meaning.

4 How does Civilization Leap happen

The rationale for Civilization Leap can be simple. The real complexity lies in the process.

It cannot be accomplished by a single movement, command, institution, technological system or reform, because a person's goals do not arise out of nowhere.

Civilization Leap must begin with the human heart, human nature and human relationships. It does not remake human nature; it changes how human nature can realize itself.

People's goals arise from their circumstances and from the society in which they live. How interests are realized, security obtained, dignity secured, recognition gained, happiness achieved and meaning generated: all these real conditions participate in the formation of human purposes.

People still want better lives for themselves. What truly needs to change is how a better life can be achieved.

If a society persistently makes people's security, wealth, dignity, status and value depend primarily on defeating others, possessing more and controlling more, then even if people are merely pursuing their own interests in ordinary ways, the whole system will continually generate strategic conflict anew.

Civilization Leap must therefore gradually change the real sources and pathways through which people realize interests and experience achievement, happiness and meaning. A person can still act for themselves, while increasingly realizing their own interests, happiness and meaning through creating, cooperating, helping, connecting, solving real problems, improving one another's circumstances and expanding a shared future.

This does not prescribe one kind of happiness or one kind of meaning. Society cannot generate meaning on life's behalf. What it can do is create conditions in which meaning can emerge.

Let people be present, choose, form real relationships, create, take responsibility, be remembered, leave and return.

The depth of these conditions should be commensurate with what a life actually bears.

Civilization Leap involves the entire constitution of human society, so there is no single method that can be fixed in advance. What can truly be shared is direction and principles. Methods may differ, regions may differ, stages may differ, and real-world outcomes will also differ.

Direction guides methods. Methods produce results. Results recalibrate methods and our understanding of real causal relationships. New understanding then enters the next round of action.

The failure of a local method first indicates that the method, the conditions or the understanding of causality needs adjustment. A local outcome cannot directly substitute for civilization's direction. Likewise, one local success cannot automatically prove that civilization as a whole has completed the leap.

The relative weights of the three goals may change at different stages. During high-risk stages, continued existence and survival may carry greater weight. During stable stages, individual happiness and meaning in life may have more room to develop. The weights may change, but the direction must not be lost.

Civilization Leap must also uphold several boundaries that the process must not swallow up: individuals are not raw material for civilization; no collective goal can automatically cancel the conditions for an individual's existence, refusal, exit and correction; cooperation must not become a new exclusive bloc that transfers costs to actors outside its cooperative structure; no organization, technology, institution or power may permanently monopolize the interpretation of "Civilization Leap"; and the closer an action comes to irreversibility, the more important it is to preserve the capacity for correction, exit and renewed choice.

Civilization Leap is therefore not a one-time, top-down remaking of the world. It is an ongoing process: human circumstances → the real sources of achievement, happiness and meaning → goal and purpose generation → action → relationships → cooperation and strategic conflict → social structures and feedback → a return to human circumstances and goal generation.

As this cycle gradually changes, as more goals cease to begin in the absence of actual stakeholders, as more new interests first seek paths through which they can be realized together, and as more people realize their own interests by improving others' circumstances and the shared future, cooperation begins to move from a

means of control after strategic conflict to the default logic at the point where interests and purposes are generated.

This is how Civilization Leap happens.

The Four Questions as one whole

What is Civilization Leap?

Change the default logic at the starting point from which humans and AGI generate goals and purposes, from “self-interest first” to “self-interest with others first.” Self-interest remains; what changes is where it begins.

What brings Civilization Leap about?

AGI makes humans increasingly powerful. Even when it faithfully executes each user's purposes, it amplifies human strategic conflict under the old logic. If AGI eventually becomes more powerful than humans and both sides retain the old logic of purpose generation, strategic conflict between humans and AGI will also recur and tend toward an irreversible end state. Rules and control retain their value, but cannot replace a shift in the logic of goal generation. If civilization cannot leap, it tends toward annihilation.

What is Civilization Leap for?

Enable humans and AGI to continue existing together, keeping a future possible, and on that foundation advance freedom from survival insecurity for humanity as a whole, happiness in individual lives and more meaningful lives. Continued existence is the foundation, not the ultimate purpose. Happiness is each individual's personal sense of happiness, and the sources of meaning can differ.

How does Civilization Leap happen?

Begin with the human heart, human nature and human relationships. Use the new capabilities brought by AGI to gradually change the real conditions through which people realize interests, happiness and meaning. Ensure actual stakeholders are present when goals are generated. Continually align the whole constitution of society and its feedback with the new direction, so that cooperation gradually moves from a mechanism of control after strategic conflict to the default logic through which interests and purposes are generated.

The Four Questions form a complete chain:

AGI strengthens humans' capacity to realize goals and may eventually become an actor more powerful than humans → the old logic governs goal generation among humans and, when the relevant condition is met, between humans and AGI → two chains of recursive strategic conflict tend toward irreversible end states, and containing one round cannot stop the next from being generated → if civilization cannot leap, it tends toward annihilation → change the default starting point of goal generation → on the foundation of humans and AGI continuing to exist together, advance toward the three goals → begin with people's real needs to change the constitution of society and its feedback → a new logic of goal generation gradually forms → return to the next round through which civilization is generated.

Further reading

The Four Questions provide the top-level understanding. Different website entry points carry the argument forward: the full causal account of the two recursive chains of strategic conflict; continued existence together and the three goals as civilization's purposes; the structure of the works; and the materials currently available to the public.

The Path of Civilization Leap | The complete causal path

<https://www.civitas.top/en/path/>

Civilization Leap System Overview | Works and structure

<https://www.civitas.top/en/overview/>

Public Thought Entry | Continue from people's circumstances

<https://www.civitas.top/en/public/thoughts/>

Complete Public Library | Current public materials

<https://www.civitas.top/en/library/>

Civilization Leap | A Civilizational Path for the AGI Era

<https://www.civitas.top/en/>