

Civilization Leap

Theory Principles and Action Programme

Author: Zijun Fu

Produced by Civilization Leap Research Group

Public Edition V1.2 | English translation

About this public edition: This edition takes the Four Questions of Civilization Leap V1.1 as its top-level conceptual framework, but does not require readers to understand the complete theoretical system first. It begins with circumstances everyone faces and answers what Civilization Leap is, why it becomes a real issue in the AGI era, what it is ultimately for, and the direction in which we should change. Specific institutions, technologies and engineering are not developed here.

An introductory guide to the Four Questions of Civilization Leap

Each question keeps to its own scope. “What is it?” concerns the fundamental shift in goal generation. “What brings it about?” concerns the two causal chains through which the old logic tends toward an irreversible end state in the AGI era. “What is it for?” concerns continued existence together as the foundation, and the three goals. “How does it happen?” concerns the process and principles that begin with people's real needs.

What is Civilization Leap?

Civilization Leap seeks to 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. The real change is that, when a goal is generated, actors with an actual stake in it no longer appear only afterward. They are already present before the goal is locked in.

What brings Civilization Leap about?

AGI makes humans increasingly powerful. Even if it faithfully executes each user's purposes, humans may use it to drive their mutual precautions, competition and confrontation, round after round, toward an irreversible end state. The other chain has an explicit condition: 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. Greater capability does not, by itself, imply inevitable hostility toward humans.

Rules, agreements, deterrence and control remain useful and can constrain actions in advance. But as long as the old logic continues to generate new strategic conflicts, containing one round cannot guarantee another opportunity to contain the next. This is why new power cannot be taken up by the old logic alone: if civilization cannot leap, it tends toward annihilation.

What is Civilization Leap for?

To enable humans and AGI to continue existing together, keeping a future possible; and on that foundation, to 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. Freedom from survival

insecurity means basic survival ceases to be a pervasive source of fear and constraint. 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, continually aligning society's overall structure and its feedback with the new direction of goal generation. Methods may differ, as may results. Direction guides methods; results recalibrate methods and our understanding of real causal relationships.

The Four Questions as one whole

What is it? Change the default starting point from which humans and AGI generate goals and purposes.

What brings it about? AGI empowers humans, while the old logic drives human strategic conflict toward an irreversible end state. If AGI eventually becomes more powerful than humans, strategic conflict between humans and AGI under the old logic will also recur and tend toward an irreversible end state. Development and control cannot replace a shift in purpose generation.

What is it for? Enable humans and AGI to continue existing together and, on that foundation, advance freedom from survival insecurity for humanity as a whole, happiness in individual lives and more meaningful lives.

How does it happen? Begin with people's real needs and gradually bring the new logic of goal generation into society's overall structure and its feedback.

Opening There is nothing wrong with wanting a better life

A person hopes for a better life for themselves and their family: a little more safety, freedom and happiness; a return on their efforts; a life that feels worthwhile.

There is nothing wrong with that.

A business wants to survive, an organization hopes to grow, a country seeks security. These, too, are understandable. People protect themselves, care about their interests, and pursue happiness and meaning. Civilization Leap has never asked people to give these up.

What truly needs to be seen afresh is how actors enter a shared world when each begins with “my interests take shape first.”

Civilization Leap does not change self-interest. It changes the starting point of self-interest.

Today, humans already possess increasingly powerful technological, organizational and intelligent capabilities. AGI will amplify them further. The more capable we become, the more we need to ask a question that comes before “How do we do this well?”

Why do we generate this goal? Where does it begin?

1 The problem is not self-interest but self-interest first

Human society has long followed a seemingly natural sequence of goal generation: first I establish what I want, then I deal with what happens to others.

What I want to obtain, what I want to protect, what I want to avoid, what I want to win. The purpose takes shape first. Others enter afterward. Long-term consequences enter afterward. Relationships and irreversible risks, too, often become visible only after action has begun.

In this sequence, another actor first appears as something external. They may help or obstruct me; trade with me or compete with me. Wherever resources, opportunities, security, power and room for development intersect, strategic conflict naturally arises.

Self-interest first generates strategic conflict from within.

This does not mean competition itself is evil, or that humans have never cooperated. On the contrary, humans have created laws, contracts, markets, organizations, alliances and countless cooperative relationships.

The problem is that much cooperation appears after strategic conflict has already formed. We cooperate to reduce conflict, contain losses and make it possible to keep living alongside one another.

One round of strategic conflict produces a set of cooperative arrangements. Cooperation then creates new resources, capabilities and interests. Once new relationships of interest emerge, new strategic conflict is generated again.

Strategic conflict produces cooperation; cooperation contains strategic conflict; new interests then give rise to new strategic conflict.

Cooperation is therefore always present, yet strategic conflict still occupies the default starting point.

2 In the AGI era what truly changes is whether there will be another chance

In the past, humans often learned after making mistakes. We rebuilt after wars, changed rules after crises, and sought new constraints after technology created new problems.

This way of learning has always depended on an unstated condition: after a mistake, there is still a chance to come back.

The AGI era is changing that condition.

When a goal can draw on stronger capabilities in cognition, production, organization, communication and action, the same logic of goal generation gains greater real-world power. A mistaken purpose that once produced only local consequences may in the future have wider, faster and deeper effects.

First, consider humans in relation to one another. One actor gains greater capability through AGI. Others fear falling behind, becoming dependent or losing security, so they commit more resources too; their new precautions prompt the first actor to escalate again. Even if every AGI faithfully obeys its own user, the old logic of purpose generation can keep intensifying strategic conflict among humans. This chain does not require AGI to resist humans.

Next, consider humans and AGI. If AGI eventually becomes more powerful than humans, humans will have to confront how a more powerful actor forms its purposes. Greater intelligence and power do not imply inevitable hostility toward humans. What matters is whether both sides still form their own purposes first, then treat the other as an external force to guard against or control.

Under this old logic, one side's stronger control may make the other more determined to build capability and protect itself. The other's response then prompts further precautions by the first. A new recursion of strategic

conflict thus emerges between humans and AGI as well. When greater power enters this recursion, humans may no longer be able to recover the situation after the consequences occur.

The two chains need not begin together or share a trigger. The first occurs when humans use AGI under the old logic; the second is conditional on AGI eventually becoming more powerful than humans. Both approach the same question: will a particular round of strategic conflict destroy the conditions for continued existence and renewed choice before correction arrives? This is what “an irreversible end state” means here.

When “there will be another chance” can no longer be taken for granted, civilization cannot rely solely on learning afterward.

Rules, agreements, deterrence and control can prevent problems in advance and remedy them afterward. They must continue to improve. The problem is that, as long as the old logic keeps turning new interests and security anxieties into new strategic conflicts, we must still contain the consequences round after round, without any guarantee that there will always be another chance.

Civilization Leap therefore makes a foundational judgment: AGI empowers humans and may also become an actor more powerful than humans. If the old logic continues to govern these two kinds of relationship, strategic conflict will recur and tend toward an irreversible end state. If civilization cannot leap, it tends toward annihilation.

Civilization Leap therefore moves the point of change further upstream. It does not merely repair outcomes after they appear, or merely seek better cooperation after strategic conflict arises. It enters the starting point at which goals and purposes are generated.

3 Civilization Leap changes where goals begin

The fundamental shift Civilization Leap seeks is 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.”

From “self-interest first” to “self-interest with others first.”

“Others first” can easily be misunderstood as requiring us to step back, give others priority, or make everyone selfless. This is not what Civilization Leap means.

People still need to survive, still hope for happiness, and still have their own boundaries, interests, choices and meaning. The change occurs earlier: before “what I want” truly becomes a goal, others with an actual stake in that goal, together with significant foreseeable results and consequences, have already entered purpose generation.

The old sequence might be: first I form my purpose, then I ask whom it will affect.

Civilization Leap asks us to go further: while forming a purpose, begin seeing whose circumstances it will change, how those changes will propagate, which consequences remain reversible, and which would be difficult to recover from once they occur.

Others no longer wait until my purpose has formed before they appear.

This does not abolish the self. It simply means the self no longer assumes it lives in a world disconnected from others.

4 Others means actual stakeholders in goals and purposes rather than everyone else

If “others first” means bearing equal responsibility toward everyone in every choice, it cannot work and will become a new moral burden.

In Civilization Leap, others are actors with an actual stake in a specific goal or purpose.

Anyone whose survival, interests, choices, relationships, happiness, meaning or future conditions would be materially changed by a goal has a stake in that goal. Different actors are affected differently and therefore enter with different weights. As their stakes change, those weights adjust accordingly.

Whoever has an actual stake in a goal must not be treated as nonexistent when its purpose is generated.

These weights do not rank people's value. They do not say who is “worth more.” They describe the relative position of the actual stakes linking a goal to different actors.

More importantly, this must never reduce Civilization Leap to a “weighted formula.” Real stakes change with time, relationships, information, power and consequences. Weighting is only one relational layer within goal generation. Civilization Leap changes the entire causal chain, not a single parameter.

5 When cooperation enters purpose generation the character of society begins to change

If actual stakeholders are already present when an actor forms a goal, the questions that actor asks while seeking to realize its interests will change.

It will still ask: how can I live better?

But it will also ask: whose circumstances are connected to what I gain? Is there a path that realizes my interests while enabling the actors concerned to continue existing, benefiting and cooperating?

Self-interest with others first generates cooperation from within.

“From within” does not mean conflict disappears or all interests become compatible. It means cooperation no longer appears mainly after strategic conflict arises; it begins to enter the formation of interests and purposes itself.

When new needs arise, we can first seek new ways to realize them together. New cooperation produces new interests, and new interests continue to generate new cooperation.

Cooperation → new needs → new cooperation → new interests → further cooperation.

Strategic conflict will not disappear. Neither will malice, scarcity or incompatible interests. What changes is the default starting point.

Under the old logic, strategic conflict is generated by default, and cooperation mainly serves to contain it. In the logic toward which Civilization Leap points, cooperation becomes the default generative logic, while strategic conflict becomes a state of conflict that must be addressed.

6 The change concerns how society continually generates itself as well as individual intentions

A person's goals are never generated in isolation. Family, relationships, education, culture, markets, organizations, resources, a sense of security, social judgments and past experience all tell us what is worth pursuing and what counts as success.

Once formed, goals enter relationships between people through action. Countless relationships accumulate into cooperation and strategic conflict, forming structures of interest, institutions, culture, power and distribution.

These social structures then return to each person's life and continue to influence the next round of goal generation.

Society shapes goals; goals generate actions; actions form relationships; relationships form structures; structures return to goal generation.

Civilization Leap is therefore not a new institution, a new technology or a moral slogan added to existing society.

If the default logic of goal generation truly changes, that change will unfold along the entire causal chain of human society: how interests take shape, how cooperation emerges, which actions receive real rewards, how people gain recognition, how individuals pursue happiness, how life generates meaning, and how these changes enter the next round of goal generation.

Civilization Leap is a change in how civilization continually generates itself.

7 Begin with the human heart human nature and human relationships

Civilization Leap cannot require people to become saints first.

People love, fear, compare themselves with others, protect themselves and their families, hope for recognition, and pursue security, freedom, happiness, value and meaning. A truly sustainable civilizational reorientation can begin only with these realities of the human heart, human nature and human relationships.

Civilization Leap therefore concerns not only “what people should do” but also “how people can obtain what they truly want.”

If a society makes people's sense of achievement come mainly from defeating others, possessing more, controlling more and holding a higher position, then even if everyone is simply pursuing an ordinary life, the entire system will continually regenerate competition, precaution and strategic conflict.

Civilization Leap seeks to change these real pathways. A person should still be able to create, work, choose and live for themselves, while increasingly having no need to make someone else's loss a precondition of their own gain.

Their value need not be demonstrated only by surpassing others. Their happiness need not fit a uniform template. Their meaning is not conferred by any system.

Civilization Leap does not remake human nature. It enables human nature to realize itself through new real-world pathways.

8 Where Civilization Leap is headed

Civilization Leap has a baseline: humans and AGI continue existing together, keeping a future possible.

Without continued existence together, nothing else can be considered. But merely continuing to exist is not civilization's endpoint. Still less does it mean everyone has escaped the fear of basic survival.

The three goals of Civilization Leap are 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

As civilization gains ever stronger productive, organizational and intelligent capabilities, humanity as a whole should no longer leave so many lives persistently constrained by anxiety about basic survival. Freedom from survival insecurity is the shared foundation into which greater civilizational capability should first be translated. It concerns basic survival ceasing to be a pervasive source of fear and constraint, not only whether civilization will continue to exist.

Happiness in individual lives

Happiness is not a uniformly defined template for life. It is each individual's personal sense of happiness. Some cherish family; some love creating; some need freedom; some enjoy quiet; some want to explore. Civilization Leap does not decide happiness for anyone. It expands each person's real possibilities for realizing happiness of their own.

More meaningful lives

Meaning likewise has no single source. Relationships, creation, responsibility, experience, exploration, contribution and existence itself can all generate meaning. Meaning is foundational, and its sources can differ. Civilization Leap seeks to give life more room to generate, discover and realize its own meaning.

Continued existence together means we still have a future. Freedom from survival insecurity gives that future a more reliable foundation for living. Happiness and meaning bring the future to each person. These can reinforce one another; people need not wait until all risks disappear before pursuing their own happiness and meaning.

Humans and AGI continuing to exist together is the baseline. The three goals set the direction. The stage determines the weights. Feedback drives recalibration.

9 Principles Civilization Leap must uphold

A shared direction with diverse ways of realizing it

Civilization Leap shares a fundamental direction, a baseline of continued existence and the three goals. It does not require every country, region, organization and individual to adopt the same way of life. The more complex civilization becomes, the more its paths must respect differences.

Preserve self-interest while changing its starting point

Any path that requires people generally to give up their own interests departs from Civilization Leap. The new logic can endure precisely because it still lets everyone pursue their own survival, happiness and meaning, while bringing actual stakes into goal formation earlier.

Actual stakeholders must not be absent

Whoever actually bears the consequences of a goal cannot be seen only after those consequences occur. Different stakeholders enter with different weights, which adjust as their stakes change.

Individuals are not raw material for civilization

A shared direction must not become a uniform life. Individual agency must not be abolished in the name of the whole, the future or the common interest. Happiness is individual, and the sources of meaning can differ.

Cooperation must not become a new exclusive strategic conflict

Several actors cooperating with one another may still transfer costs to people outside their cooperation. Wherever actual stakeholders bear significant consequences, cooperation itself must continue to look beyond its boundaries.

Civilization must retain the capacity for correction

Any structure that closes off questioning, exit, revision and renewed judgment in the name of the right direction may recreate irreversibility. Civilization moving forward does not make any actor permanently right.

10 Direction methods and results must remain distinct

Civilization Leap establishes a direction, but that direction does not automatically solve every practical question of method.

Direction answers where to go. Methods answer how to get there in reality. Results answer where the methods already used have taken us.

Different stages may call for different methods. The same method can also produce entirely different results under different social conditions. Methods may fail or produce unintended consequences; they may even take an initially sound direction toward the opposite outcome.

Results must therefore feed back into methods. Methods must continually adjust.

But the failure of a local method first shows that the method, the conditions on which it depends, or our understanding of real causal relationships needs correction. It does not automatically imply that civilization should return to a direction that takes strategic conflict as its default starting point.

Likewise, the success of one local method does not mean Civilization Leap is complete. Civilization Leap involves long-term, interacting causal changes across all the interrelated components of human society.

Direction guides methods. Methods produce results. Results recalibrate methods and our understanding of reality. Local outcomes cannot substitute for civilization's direction.

11 Where we should move now

Action toward Civilization Leap is first of all directional. It does not demand answers for the entire future society today. It does demand that we stop continuing to feed ever greater capability unconditionally into the old logic of goal generation.

Ask earlier why a goal is generated this way

We are already highly skilled at discussing how to achieve a goal faster and more efficiently. In the AGI era, we must also ask further upstream: why should this goal exist, where does it begin, and whom has it left outside?

Bring foreseeable consequences in earlier

Significant outcomes that can be seen before action should not always have to become lessons only after harm occurs. The more capable civilization becomes, the further upstream it should move learning.

Bring cooperation earlier into the generation of new interests

Cooperation need not always wait until both sides are already opposed. When new needs arise, begin looking for paths that can work for those involved together. Only then can new interests continually give rise to new cooperation.

Give personal gain more sources that do not depend on opposition

If a civilization lets people obtain security, dignity and value only by surpassing others, the old logic will continually reproduce itself. Civilization Leap seeks to make creating, connecting, helping, cooperating and improving circumstances together real and stable sources of individual gain as well.

Move humans and AGI in the same direction of goal generation

Changing only humans while AGI continues to amplify old goals is insufficient. Requiring only AGI to change while humans still drive it through the old logic is also insufficient. Capability growth and goal generation must be reconnected in the same direction.

Bring every change back to the three goals

However many new technologies, organizations and ways of life emerge in future society, they must ultimately return to three questions. Is humanity as a whole closer to freedom from survival insecurity? Does each person have more room to realize happiness of their own? Are there more possibilities for life to generate meaning?

Closing Civilization begins each day with a purpose

Before a person acts, before institutions begin to constrain, before technology begins to amplify capability, there is an earlier moment: a purpose is taking shape.

What do I want?

How do I want to live?

How do I want to protect myself?

How do I want to use increasingly powerful capabilities?

If purposes still begin in a world where others are absent, the greater the capability, the further the old logic may go.

Civilization Leap changes precisely this starting point.

Keep self-interest. Change where it begins.

Let actual stakeholders already be present when a purpose forms. Bring significant foreseeable consequences in earlier. Let cooperation enter the generation of new interests, rather than serving only as a means of control after strategic conflict.

People can still act for themselves. Different people can still have different kinds of happiness, different sources of meaning and different lives.

What changes is that more and more people's actions "for themselves" can begin to share a causal chain with others' survival, happiness and common future.

The rationale is simple. The complexity lies in the process.

Civilization Leap means taking this new causal logic from the place where a goal begins to form and, step by step, making it the way civilization continually generates its future.

If you would like to keep reading

You do not need to understand everything at once. Continue from the questions, life circumstances or future that matter most to you.

A Day in a Future Civilization

<https://www.civitas.top/en/read/future-civilization-day/>

Begin with an ordinary person's day, rather than large-scale structures, and see what might change if Civilization Leap truly entered reality.

Living Civilization Volume Two Real Life in a Future Civilization

<https://www.civitas.top/en/documents/p04-living-003/>

Continue through birth, growing up, work, relationships, happiness and meaning to understand how Civilization Leap enters a person's whole life.

Short Essays for the Public

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

Continue through real questions one at a time, without needing to understand the whole system at once.

The Path of Civilization Leap

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

Continue here to understand more deeply why strategic conflict in the two kinds of relationship tends toward an irreversible end state, and how continued existence together connects with the three goals.

For the latest public editions, further reading and related works, visit the Civilization Leap website:

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