AI / ABUNDANCE / HUMAN PURPOSE

If AI Can Change Everything, Why Not Start With What Humanity Needs?

The promise of superintelligence demands more than a wager on utopia. It demands a serious commitment to a better life for everyone.

By Junior Williams | Opinion | September 10, 2026

What would persuade you that the race to build superintelligence was going well? A model that could improve its own code? A company pulling ahead of its competitors? A country claiming the lead?

I would like to hear more about whether people will have enough to eat, somewhere safe to live, and the freedom to spend their lives doing something other than trying to afford the next month.

Those ambitions can sound strangely modest beside promises of machines that will transform civilization. Yet they are the ambitions against which I want to judge that transformation. If AI is going to change everything, the essentials of a decent human life deserve a central place in the plan.

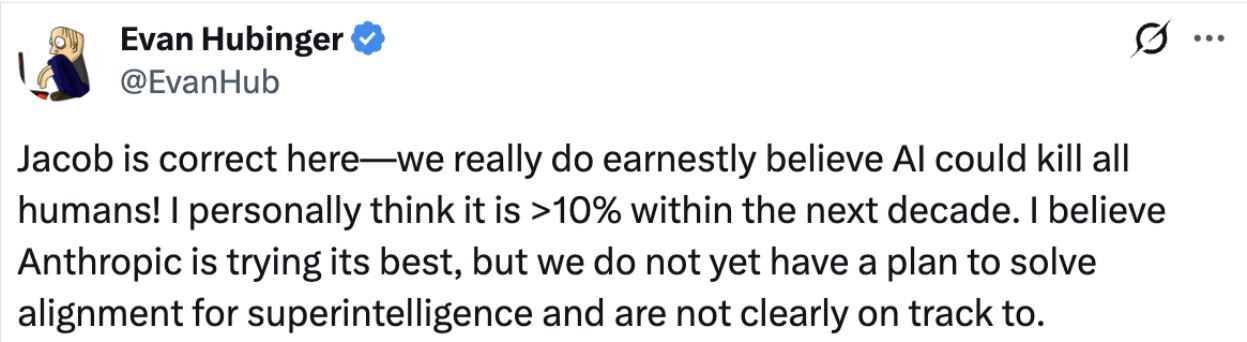
Jacob Coxon's resignation from Anthropic brought the question into focus. In his September 8 post, he accused Anthropic and OpenAI of racing toward self-improving superintelligence and gambling with people's lives. That is his assessment, and a serious one. The disagreement that followed deserves attention. But it also leaves room for a question that an argument over extinction probabilities cannot answer: what are we collectively trying to build? [1]

My answer begins with food, clothing, shelter and education. I regard these as universal rights. I want AI directed toward making them reliably available to everyone, alongside tackling climate change and expanding human freedom. **That is a minimum obligation.** It does not make unjustified risks acceptable. Promising a worthy destination cannot settle whether the journey is safe enough to take.

Jacob Coxon's resignation post. His warning prompted the discussion examined here. Screenshot captured September 10, 2026.

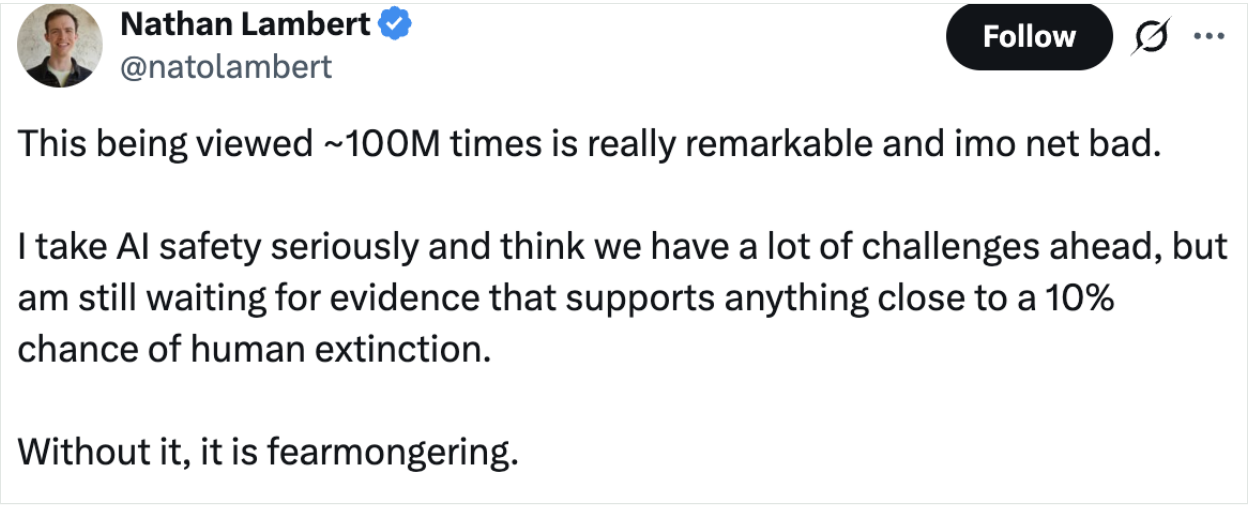
The debate leaves a question unanswered

Evan Hubinger's response was striking. The Anthropic alignment researcher offered a personal estimate of more than a 10 percent chance of human extinction within the next decade. In a subsequent clarification, he distinguished his concerns about future recursive self-improvement from what he considered the low risk of present models. The distinction matters: this was a personal forecast about a possible trajectory, not a measured probability or an official company assessment. [2]



Evan Hubinger's personal assessment. His later clarification concerns the difference between present models and future recursive self-improvement.

AI researcher Nathan Lambert challenged the evidence for estimates of that magnitude. His response illustrates why a widely shared warning should be examined rather than accepted on the strength of its reach. Expertise deserves a hearing; competing judgments still need reasons. [3]



Nathan Lambert's response, questioning the evidentiary basis for the extinction estimate.

I cannot adjudicate those forecasts from an X thread. I also do not think the public should have to settle them before asking what obligations accompany this extraordinary project. Safety and purpose are separate questions, both requiring an answer. A system could be safer than feared and still deliver a profoundly disappointing future.

Imagine that the most dramatic forecasts of abundance prove broadly right, but access to essentials remains precarious for people without sufficient income. We might have achieved astonishing technical progress while leaving a basic human problem intact. Calling that outcome a success would require a very narrow definition of success.

That is what troubles me about a debate whose horizon can stretch from catastrophe to utopia while the route to an ordinary, secure life remains vague. I want more attention paid to that route. It should be possible to demand greater seriousness about both the risks and the benefits without joining a camp that asks us to dismiss either.

Take the promise seriously

It would be unfair to suggest that frontier labs have never considered these questions. In *Machines of Loving Grace*, Dario Amodei discusses poverty, economic development, health, governance, and the meaning of work. He also acknowledges practical limits on what greater intelligence can accomplish. OpenAI's *Industrial policy for the Intelligence Age* presents exploratory proposals intended to spread prosperity and invites public debate. These are substantive starting points. [4,5]

My challenge is to take such ambitions seriously enough to ask how they become commitments. Who receives the benefits? What resources are attached? How would we know whether an initiative was improving lives, and what would change if it failed?

A company should have room to experiment, make mistakes and discover uses nobody anticipated. Basic research cannot be reduced to a quarterly social-impact target. But the freedom to explore does not prevent a company from making explicit commitments about what it intends to help society achieve. Commercial uncertainty is familiar territory for these organizations. Human benefit deserves comparable intellectual effort.

The strongest invitation to the industry is to bring its ingenuity to that task. Treat the difficulty of delivering abundance as a problem worthy of exceptional people. Give the relationship between technical capability and public benefit the attention currently given to making the technology more capable.

An impressive demonstration can begin that work. The harder achievement is a result that survives outside the demonstration: a service that people can afford, that local institutions can maintain, and that continues working after the announcement has passed.

Start with the life we want

The future I keep returning to resembles the social aspiration of *Star Trek*: material security, room for curiosity, and a human identity larger than the country of one's birth. Its appeal is the freedom to learn, explore and contribute when survival no longer depends on a paycheque. The franchise's own discussion of its economic vision makes that connection between abundance and individual freedom explicit. [6]

I use that fictional destination to ask a practical question. What would we need to start doing now to move in that direction? We can work backwards without pretending that replicators are an engineering plan or that a television series resolves political disagreements.

For me, the starting point is straightforward. *Access to a decent life should not depend on whether the labour market currently values what a person can sell.* People should be able to care for relatives, study, create, conduct research or build something new without their basic needs becoming collateral for the decision.

That is a demanding ambition. It would require choices about resources, institutions and responsibilities that a more intelligent machine cannot legitimately make on humanity's behalf. Nor would plenty automatically eliminate crime, cruelty or conflict. A society with less desperation would still need rights, trust and ways to resolve disputes. My hope is for the freedom and security that make a more peaceful life possible.

An industry comfortable imagining radical changes in intelligence should be willing to entertain equally serious changes in what people can expect from life.

Make abundance tangible

Food is a useful place to begin. An AI initiative could investigate better crop planning, reduced spoilage or more reliable distribution. Those are possible avenues, not a claim that one model can end hunger. A meaningful project would need farmers, logistics operators and local communities to determine where assistance is useful. Its success would depend on whether people actually became more food-secure.

Housing invites the same discipline. We could direct technical effort toward construction methods, materials, energy efficiency and planning. But a cheaper design does not necessarily become an affordable home. Land, financing, infrastructure and the way savings are passed on all matter. The goal should reach as far as the household receiving the keys.

In education, imagine assistance shaped around a learner's language, circumstances and goals, with teachers involved in its design and evaluation. Ask whether understanding improves, whether errors are caught, and whether the service remains accessible to families with little money. Access to a chatbot is a beginning, not an education in itself.

Efficient transportation and climate action require similar care. Better planning or energy management could be worth exploring, but the test should include what happens in the physical world. Are journeys more reliable? Does a proposed improvement reduce emissions once its own resource requirements are counted? Can the community operate it without permanent dependence on an expensive outside provider?

These examples reveal how much of the work lies beyond a model. Local knowledge, physical investment, sound institutions and implementation all belong in the picture. That should expand the mission rather than excuse the technology from it. We need collaboration between those building AI and those who understand the problems it is supposed to help solve.

We should also be willing to discover that AI is not the best intervention in a particular case. If an ordinary infrastructure improvement would do more good, a credible programme would act on that finding. The objective is better lives. The technology earns its place by helping deliver them.

There is also a difference between making something available and making access secure. A family cannot build its future around a benefit that disappears when a pilot ends or a company changes strategy. If we take basic needs seriously, continuity belongs among the design questions from the beginning. Who maintains the service? Who can challenge an exclusion? What happens when the provider leaves? These questions may sound mundane beside superintelligence. For the person relying on the result, they are decisive.

Freedom is an economic ambition

A committed capitalist might reasonably ask who will pay for this, what keeps innovation attractive, and who prevents a public programme from wasting resources. Those questions deserve answers. Entrepreneurs and investors take risks, and a system that stops rewarding productive effort could undermine the capabilities we hope to develop.

But a defence of enterprise should also ask what its success makes possible. If increasingly capable machines can do more of the work needed to support us, why should access to the essentials remain tied indefinitely to everyone finding a buyer for their labour?

There is an argument here that takes individual freedom seriously. A person with secure housing and enough to eat may have more room to start a business, retrain, leave an exploitative situation or attempt work that does not pay immediately. Material security could widen the pool of people able to take a productive risk. It need not imply that ambition disappears.

I am not proposing that every difference in reward be abolished. I am questioning the necessity of deprivation as the price of an economy that rewards achievement. We can value invention while asking whether the gains from extraordinary productivity should establish a more generous minimum for everyone.

The transition would be difficult. Resources are finite, even when a software demonstration makes abundance feel close. Important work still has to be done, and people will disagree about a fair contribution. Those are reasons to examine the idea carefully. They do not establish that dependence on employment must remain the permanent foundation of human security.

We should think about that transition before people experience it as a personal failure. If a society deliberately develops systems that can perform more human work, it ought to consider how people will retain security, standing and choice when particular skills become less valuable. Telling everyone to adapt is incomplete unless we also ask what they are being offered a fair opportunity to adapt to. Some will want another career. Others may want time to study or care for someone. A humane transition should make room for both.

I would like advocates of markets to help develop the strongest version of this argument. What would an economy look like if it preserved competition and choice while making destitution an unacceptable outcome? AI gives that longstanding question a new urgency, even though it does not answer it for us.

Humanity is larger than the home market

One reply to Coxon's thread reduced the strategic objection to a short question: "Do you want China to win?" Ben Pouladian's response captures the national competition framing that this discussion must confront. It does not settle what winning should mean. [7]

Ben Pouladian's reply to the final part of Coxon's thread. It illustrates the national competition objection.

Governments have responsibilities to their citizens, and security concerns cannot simply be wished away. But a claim to benefit humanity must extend beyond the countries hosting the leading laboratories. A child's entitlement to food or education should not turn on whether their country is an important customer, investor or strategic ally.

Global benefit also means giving affected people a say. A community should be able to identify the problem, challenge the proposed intervention and assess the result. Being handed a tool designed elsewhere is different from having meaningful influence over its purpose. Participation has to reach further than a launch photograph or a consultation after the major decisions are made.

The longer-term ambition, for me, reaches beyond borders themselves. I would like humanity to work toward a world where birthplace no longer determines the range of lives a person is allowed to pursue. That is part of working backwards from the future I find compelling, rather than assuming today's political arrangements are the final form of human cooperation.

Moving beyond those arrangements would raise profound questions. We can begin by asking whether our most advanced technology should help us cooperate across divisions or simply give those divisions more power. An AI race can be nationally successful and still fall short of its claimed human purpose.

An obligation worth discussing

If we are considering legal obligations for frontier AI companies, helping advance these global initiatives deserves a place in the discussion. I would support a proportionate obligation even where fulfilling it reduced potential profits. **A commitment that applies only when it costs nothing is a fragile basis for a universal promise.**

One possible starting point would be funded plans addressing basic needs, developed with affected countries and communities. They could identify resources, partners, milestones and a means of independently assessing progress. The details should be contestable. We would need to distinguish effort from results and honest failure from an initiative designed chiefly to satisfy a reporting requirement.

That would not transfer the responsibility for poverty or public services to a handful of technology companies. Governments, researchers, civil society and existing providers would remain essential. Nor should a rule become so burdensome that only the largest firms could comply. Competition, accountability and the ability to revise a failed approach would all matter.

A useful test would be whether a programme can explain its contribution without relying on a sweeping claim about the future. What improved, for whom, and compared with what would otherwise have happened? If the gains reached only people already well served, could it acknowledge that limitation? If the tool failed, would funding follow the evidence toward another approach?

It would also matter who can hold a programme to account. A donor, a government and a participating household might judge the same intervention differently. An evaluation should have room for those disagreements, including the people who found it inaccessible or harmful. A favourable average cannot tell the whole story. The purpose of scrutiny would be to learn whether the effort deserves expansion, redesign or replacement, rather than to protect a commitment once announced.

Ownership, financing and international enforcement would all need serious debate. I do think the people capable of building these systems, together with the people expected to live with their consequences, should be working seriously on them.

We can ask for more

There is a temptation to treat any positive vision as a justification for accepting the danger required to reach it. I reject that bargain. An attractive account of universal abundance does not establish that a particular development path is responsible. Safety needs its own evidence and justification, regardless of how inspiring the destination sounds.

There is an equal temptation to let the disagreement over danger consume the entire conversation. Then, if the worst does not happen, we risk accepting almost any remaining outcome as a victory. Surviving a technological transition would be an inadequate measure of how well we used it.

I want the people building AI to succeed at something larger than surpassing the previous model. I want their success to mean fewer people living in deprivation, more people able to learn and contribute, and a planet on which those gains can last. That ambition leaves plenty of room for exceptional companies and rewarding work.

That is a challenge I would welcome the industry taking up in public. Show us where the obstacles are, which responsibilities belong elsewhere, and what cooperation would make a difference. Invite criticism early enough to change the work. A candid account of a difficult effort would be more persuasive to me than certainty about a distant utopia.

We should be able to look beyond a lab's technical achievements and see a serious effort to connect them to those outcomes. We should be able to question the effort without being treated as opponents of progress. And we should be honest when a promise outruns what anyone knows how to deliver.

I do not have the blueprint for a world of abundance and freedom. I am unwilling to accept that we need one before we can ask more of this race. If AI can change everything, helping humanity secure the essentials of a decent life is a reasonable place to begin.

We need to do better.

Sources and editorial notes

Selected sources, not a complete survey of replies. Native screenshots were captured September 10, 2026; displayed counts reflect that capture. Forecasts are personal judgments. The social vision and policy possibilities are the author's opinions, not claims about current capabilities or enacted law.

1. **Jacob Coxon.** Resignation post and thread.
2. **Evan Hubinger.** Personal risk estimate and clarification about present models.
3. **Nathan Lambert.** Initial response and subsequent qualification.
4. **Dario Amodei.** Machines of Loving Grace, October 2024.
5. **OpenAI.** Industrial policy for the Intelligence Age. Exploratory proposals.
6. **Star Trek.** Jason M. Barr, Smith, Marx, and... Picard?: Star Trek and Our Economic Future, March 11, 2020. Fictional reference, not a policy template.
7. **Ben Pouladian.** Reply to the final part of Coxon's thread. Example of the strategic objection.