

# My techno-optimism

Special thanks to Morgan Beller, Juan Benet, Eli Dourado, Karl Floersch, Sriram Krishnan, Nate Soares, Jaan Tallinn, Vincent Weisser, Balvi volunteers and others for feedback and review.

Last month, Marc Andreessen published his ["techno-optimist manifesto"](#), arguing for a renewed enthusiasm about technology, and for markets and capitalism as a means of building that technology and propelling humanity toward a much brighter future. The manifesto unambiguously rejects what it describes as an ideology of stagnation, that fears advancements and prioritizes preserving the world as it exists today. This manifesto has received a lot of attention, including response articles from [Noah Smith](#), [Robin Hanson](#), [Joshua Gans](#) (more positive), and [Dave Karpf](#), [Luca Ropek](#), [Ezra Klein](#) (more negative) and many others. Not connected to this manifesto, but along similar themes, are James Pethokoukis's ["The Conservative Futurist"](#) and Palladium's ["It's Time To Build for Good"](#). This month, we saw a similar debate enacted

through the [OpenAI dispute](#), which involved many discussions centering around the dangers of superintelligent AI and the possibility that OpenAI is moving too fast.

My own feelings about techno-optimism are warm, but nuanced. I believe in a future that is vastly brighter than the present thanks to radically transformative technology, and I believe in humans and humanity. I reject the mentality that the best we should try to do is to keep the world roughly the same as today but with less greed and more public healthcare. However, I think that not just magnitude but also direction matters. There are certain types of technology that much more reliably make the world better than other types of technology. There are certain types of technology that could, if developed, mitigate the negative impacts of other

types of technology. The world over-indexes on some directions of tech development, and under-indexes on others. We need active human intention to choose the directions that we want, as the formula of "maximize profit" will not arrive at them automatically.

In this post, I will talk about what techno-optimism means to me

. This includes the broader worldview that motivates my work on certain types of blockchain and cryptography applications and social technology, as well as other areas of science in which I have expressed an interest. But perspectives on this broader question also have implications for AI, and for many other fields. Our rapid advances in technology are likely going to be the most important social issue in the twenty first century, and so it's important to think about them carefully.

## Table of contents

- [Technology is amazing, and there are very high costs to delaying it](#)
- [The environment, and the importance of coordinated intention](#)
- [The environment, and the importance of coordinated intention](#)
- [AI is fundamentally different from other tech, and it is worth being uniquely careful](#)
- [Existential risk is a big deal](#)
- [Even if we survive, is a superintelligent AI future a world we want to live in?](#)
- [The sky is near, the emperor is everywhere](#)
- [Existential risk is a big deal](#)
- [Even if we survive, is a superintelligent AI future a world we want to live in?](#)
- [The sky is near, the emperor is everywhere](#)
- [d/acc: Defensive \(or decentralization, or differential\) acceleration](#)
- [Defense-favoring worlds help healthy and democratic governance thrive](#)
- [Macro physical defense](#)
- [Micro physical defense \(aka bio\)](#)
- [Cyber defense, blockchains and cryptography](#)
- [Info defense](#)
- [Social technology beyond the "defense" framing](#)

- [Defense-favoring worlds help healthy and democratic governance thrive](#)
- [Macro physical defense](#)
- [Micro physical defense \(aka bio\)](#)
- [Cyber defense, blockchains and cryptography](#)
- [Info defense](#)
- [Social technology beyond the "defense" framing](#)
- [So what are the paths forward for superintelligence?](#)
- [A happy path: merge with the AIs?](#)
- [A happy path: merge with the AIs?](#)
- [Is d/acc compatible with your existing philosophy?](#)
- [We are the brightest star](#)
- [The environment, and the importance of coordinated intention](#)
- [Existential risk is a big deal](#)
- [Even if we survive, is a superintelligent AI future a world we want to live in?](#)
- [The sky is near, the emperor is everywhere](#)
- [Defense-favoring worlds help healthy and democratic governance thrive](#)
- [Macro physical defense](#)
- [Micro physical defense \(aka bio\)](#)
- [Cyber defense, blockchains and cryptography](#)
- [Info defense](#)
- [Social technology beyond the "defense" framing](#)
- [A happy path: merge with the AIs?](#)

## Technology is amazing, and there are very high costs to delaying it

In some circles, it is common to downplay the benefits of technology, and see it primarily as a source of dystopia and risk. For the last half century, this often stemmed either from environmental concerns, or from concerns that the benefits will accrue only to the rich, who will entrench their power over the poor. More recently, I have also started to see libertarians

becoming worried about some technologies, out of fear that the tech will lead to centralization of power. This month, I did [some polls](#) asking the following question: if a technology had to

be restricted, because it was too dangerous to be set free for anyone to use, would they prefer it be monopolized or delayed by ten years? I was surprised to see, across three platforms and three choices for who the monopolist would be, a uniform overwhelming vote for a delay.

And so at times I worry that we have overcorrected, and many people miss the opposite side of the argument: that the [benefits of technology](#) are [really friggin massive](#)

, on those axes where we can measure it the good massively outshines the bad, and the costs of even a decade of delay are incredibly high

.

To give one concrete example, let's look at a life expectancy chart:

[

]([https://ourworldindata.org/grapher/life-expectancy?time=1900..latest&country=JPN~FRA~CHN~USA~ARG~IND~DEU~RUS~OWID\\_AFR](https://ourworldindata.org/grapher/life-expectancy?time=1900..latest&country=JPN~FRA~CHN~USA~ARG~IND~DEU~RUS~OWID_AFR))

What do we see? Over the last century, truly massive progress. This is true across the entire world, both the historically

wealthy and dominant regions and the poor and exploited regions.

Some blame technology for creating or exacerbating calamities such as totalitarianism and wars. In fact, we can see the deaths caused by the wars on the charts: one in the 1910s (WW1), and one in the 1940s (WW2). If you look carefully, The Spanish Flu, the Great Leap Forward, and other non-military tragedies are also visible. But there is one thing that the chart makes clear: even calamities as horrifying as those are overwhelmed by the sheer magnitude of the unending march of improvements in [food](#), [sanitation](#), [medicine](#) and infrastructure that took place over that century.

This is mirrored by large improvements to our everyday lives. Thanks to the internet [most people around the world](#) have access to information at their fingertips that would have been unobtainable twenty years ago. The global economy is becoming more accessible thanks to improvements in international payments and finance. Global poverty is [rapidly dropping](#). Thanks to online maps, we no longer have to worry about getting lost in the city, and if you need to get back home quickly, we now have far easier ways to call a car to do so. Our property becoming digitized, and our [physical goods becoming cheap](#), means that we have much less to fear from physical theft. Online shopping has reduced the disparity in access to goods between the global megacities and the rest of the world. In all kinds of ways, automation has brought us the eternally-underrated benefit of simply making [our lives more convenient](#).

These improvements, both quantifiable and unquantifiable, are large

. And in the twenty first century, there's a good chance that even larger improvements are soon to come. Today, ending aging and disease seem utopian. But from the point of view of [computers as they existed in 1945](#), the modern era of putting chips into pretty much everything would have seemed utopian: even science fiction movies often kept their computers room-sized. If biotech advances as much over the next 75 years as computers advanced over the last 75 years, the future may be more impressive than almost anyone's expectations.

Meanwhile, arguments expressing skepticism about progress have often gone to dark places. Even medical textbooks, like this one in the 1990s (credit [Emma Szewczak](#) for finding it), sometimes make extreme claims denying the value of two centuries of medical science and even arguing that it's not obviously good to save human lives:

The "[limits to growth](#)" thesis, an idea advanced in the 1970s arguing that growing population and industry would eventually deplete Earth's limited resources, ended up inspiring [China's one child policy](#) and [massive forced sterilizations in India](#). In earlier eras, concerns about overpopulation were used to [justify](#) mass [murder](#). And those ideas, argued [since 1798](#), have a long history of [being proven wrong](#).

It is for reasons like these that, as a starting point, I find myself very uneasy about arguments to slow down technology or human progress. Given how much all the sectors are interconnected, even sectoral

slowdowns are risky. And so when I write things like what I will say later in this post, departing from open enthusiasm for progress-no-matter-what-its-form, those are statements that I make with a heavy heart - and yet, the 21st century is different and unique enough that these nuances are worth considering.

That said, there is one important point of nuance to be made on the broader picture, particularly when we move past "technology as a whole is good" and get to the topic of "which specific technologies are good?". And here we need to get to many people's issue of main concern: the environment.

## The environment, and the importance of coordinated intention

A [major exception](#) to the trend of pretty much everything getting better over the last hundred years is climate change:

[

](<https://digg.com/2020/global-mean-temperature-past-2000-years>)

Even pessimistic scenarios of ongoing temperature rises would not come anywhere near causing the literal extinction of humanity. But such scenarios could plausibly kill more people than major wars, and severely harm people's health and livelihoods in the regions where people are already struggling the most. [A Swiss Re institute study](#) suggests that a worst-case climate change scenario might lower the world's poorest countries' GDP by as much as 25%. [This study](#) suggests that life spans in rural India might be a decade lower than they otherwise would be, and studies like [this one](#) and [this one](#) suggest that climate change could cause a hundred million excess deaths by the end of the century.

These problems are a big deal. My answer to why I am optimistic about our ability to overcome these challenges is twofold. First, after decades of hype and wishful thinking, [solar power](#) is [finally](#) turning [a corner](#), and [supportive technologies like batteries](#) are making similar progress. Second, we can look at humanity's track record in solving previous environmental problems. Take, for example, air pollution. Meet the dystopia of the past: the Great Smog of London, 1952.

[

](<https://www.britannica.com/event/Great-Smog-of-London>)

What happened since then? Let's ask Our World In Data again:

[

](<https://ourworldindata.org/london-air-pollution>)

As it turns out, 1952 was not even the peak: in the late 19th century, even higher concentrations of air pollutants were just accepted and normal

. Since then, we've seen a century of ongoing and rapid declines. I got to personally experience the tail end of this in my visits to China: in 2014, high levels of smog in the air, [estimated to reduce life expectancy by over five years](#) were normal, but by 2020, the air often seemed as clean as many Western cities. This is not our only success story. In many parts of the world, [forest areas are increasing](#). The acid rain crisis [is improving](#). The [ozone layer has been recovering](#) for decades.

To me, the moral of the story is this. Often, it really is the case that version N of our civilization's technology causes a problem, and version N+1 fixes it. However, this does not happen automatically, and requires intentional human effort

. The ozone layer is recovering because, [through international agreements like the Montreal Protocol, we made it recover](#). Air pollution is improving because we made it improve. And similarly, solar panels have not gotten massively better because it was a preordained part of the energy tech tree; solar panels have gotten massively better because decades of awareness of the importance of solving climate change have motivated both engineers to work on the problem, and companies and governments to fund their research. It is intentional action, coordinated through public discourse and culture shaping the perspectives of governments, scientists, philanthropists and businesses, and not an [inexorable "techno-capital machine"](#), that had solved these problems

## AI is fundamentally different from other tech, and it is worth being uniquely careful

A lot of the dismissive takes I have seen about AI come from the perspective that it is "just another technology": something that is in the same general class of thing as social media, encryption, contraception, telephones, airplanes, guns, the printing press, and the wheel. These things are clearly very socially consequential. They are not just isolated improvements to the well-being of individuals: they radically transform culture, change balances of power, and harm people who heavily depended on the previous order. Many [opposed them](#). And on balance, the pessimists have invariably turned out wrong.

But there is a different way to think about what AI is: it's a new type of mind

that is rapidly gaining in intelligence, and it stands a serious chance of overtaking humans' mental faculties and becoming the new apex species on the planet. The class of things in that

category is much smaller: we might plausibly include humans surpassing monkeys, multicellular life surpassing unicellular life, the [origin of life itself](#), and perhaps the Industrial Revolution, in which machine edged out man in physical

strength. Suddenly, it feels like we are walking on much less well-trodden ground.

## Existential risk is a big deal

One way in which AI gone wrong could make the world worse is [almost](#) the worst possible way: it could [literally cause human extinction](#). This is an extreme claim: as much harm as the worst-case scenario of climate change, or an artificial pandemic or a nuclear war, might cause, there are many islands of civilization that would remain intact to pick up the pieces. But a superintelligent AI, if it decides to turn against us, may well leave no survivors, and end humanity for good. Even [Mars](#) may not be safe.

A big reason to be worried centers around [instrumental convergence](#): for a very wide class of goals that a superintelligent entity could have, two very natural intermediate steps that the AI could take to better achieve those goals are (i) consuming resources, and (ii) ensuring its safety. The Earth contains lots of resources, and humans are a [predictable threat](#) to such an entity's safety. We could

try to give the AI an explicit goal of loving and protecting humans, but we have [no idea](#) how to [actually](#) do [that](#) in a way that won't completely break down as soon as the AI [encounters an unexpected situation](#). Ergo, we have a problem.

[MIRI researcher Rob Bensinger's attempt](#) at illustrating different people's estimates of the probability that AI will either kill everyone or do something almost as bad. Many of the positions are rough approximations based on people's public statements, but many others have publicly given their precise estimates; quite a few have a "probability of doom" over 25%.

A [survey of machine learning researchers](#) from 2022 showed that on average, researchers think that there is a 5-10% chance that AI will literally kill us all: about the same probability as the statistically expected chance that [you will die of non-biological causes like injuries](#).

This is all a speculative hypothesis, and we should all be wary of speculative hypotheses that involve complex multi-step stories. However, these arguments have survived over a decade of scrutiny, and so, it seems worth worrying at least a little

bit. But even if you're not worried about literal extinction, there are other reasons to be scared as well.

## Even if we survive, is a superintelligent AI future a world we want to live in?

A lot of modern science fiction is dystopian, and paints AI in a bad light. Even non-science-fiction attempts to identify possible AI futures often give [quite unappealing answers](#). And so I went around and asked the question: what is a depiction, whether science fiction or otherwise, of a future that contains superintelligent AI that we would want

to live in. The answer that came back by far the most often is Iain Banks's [Culture series](#).

The Culture series features a far-future interstellar civilization primarily occupied by two kinds of actors: regular humans, and superintelligent AIs called Minds. Humans have been augmented, but only slightly: medical technology theoretically allows humans to live indefinitely, but most choose to live only for around 400 years, seemingly because they get bored of life at that point.

From a superficial perspective, life as a human seems to be good: it's comfortable, health issues are taken care of, there is a wide variety of options for entertainment, and there is a positive and synergistic relationship between humans and Minds. When we look deeper, however, there is a problem: it seems like the Minds are completely in charge, and humans' only role in the stories is to act as pawns of Minds, performing tasks on their behalf.

Quoting from [Gavin Leech's "Against the Culture"](#):

The humans are not the protagonists. Even when the books seem to have a human protagonist, doing large serious things, they are actually the agent of an AI. (Zakalwe is one of the only exceptions, because he can do immoral things the Minds don't want to.) "The Minds in the Culture don't need the humans, and yet the humans need to be needed." (I think only a small number of humans need to be needed - or, only a small number of them need it enough to forgo the many comforts. Most people do not live on this scale. It's still a fine critique.)

The projects the humans take on risk inauthenticity. Almost anything they do, a machine could do better. What can you do? You can order the Mind to not catch you if you fall from the cliff you're climbing-just-because; you can delete the backups of your mind so that you are actually risking. You can also just leave the Culture and rejoin some old-fashioned, unfree "strongly evaluative" civ. The alternative is to evangelise freedom by joining Contact.

I would argue that even the "meaningful" roles that humans are given in the Culture series are a stretch; I asked ChatGPT (who else?) why humans are given the roles that they are given, instead of Minds doing everything completely by themselves, and I personally found [its answers](#) quite underwhelming. It seems very hard to have a "friendly" superintelligent-AI-dominated world where humans are anything other than pets.

The world I don't want to see.

Many other scifi series posit a world where superintelligent AIs exist, but take orders

from (unenanced) biological human masters. Star Trek is a good example, showing a vision of harmony between the starships with their [AI "computers"](#) (and [Data](#)) and their human operators crewmembers. However, this feels like an incredibly unstable equilibrium. The world of Star Trek appears idyllic in the moment, but it's hard to imagine its vision of human-AI relations as anything but a transition stage a decade before starships become entirely

computer-controlled, and can stop bothering with large hallways, artificial gravity and climate control.

A human giving orders to a superintelligent machine would be far less intelligent than the machine, and it would have access to less information. In a universe that has any degree of competition, the civilizations where humans take a back seat would outperform those where humans stubbornly insist on control. Furthermore, the computers themselves

may wrest control. To see why, imagine that you are legally a literal slave of an eight year old child. If you could talk with the child for a long time, do you think you could convince the child to sign a piece of paper setting you free? I have not run this experiment, but my instinctive answer is a strong yes. And so all in all, humans becoming pets seems like an attractor that is very hard to escape.

## The sky is near, the emperor is everywhere

The Chinese proverb 天高皇帝远 ("tian gao huang di yuan"), "the sky is high, the emperor is far away", encapsulates a basic fact about the limits of centralization in politics. Even in a nominally large and despotic empire - in fact, especially

if the despotic empire is large, there are practical limits to the leadership's reach and attention, the leadership's need to delegate to local agents to enforce its will dilutes its ability to enforce its intentions, and so there are always places where a certain degree of practical freedom reigns. Sometimes, this can have downsides: the absence of a faraway power enforcing uniform principles and laws can create space for local hegemony to steal and oppress. But if the centralized power goes bad, practical limitations of attention and distance can create practical limits to how bad it can get.

With AI, no longer. In the twentieth century, modern transportation technology made limitations of distance a much weaker constraint on centralized power than before; the great totalitarian empires of the 1940s were in part a result. In the twenty



first, scalable information gathering and automation may mean that attention will no longer be a constraint either. The consequences of natural limits to government disappearing entirely

could be dire.

Digital authoritarianism has been [on the rise for a decade](#), and surveillance technology has already given authoritarian governments powerful new strategies to crack down on opposition: let the protests happen, but then detect and [quietly go after](#) the participants [after the fact](#). More generally, my basic fear is that the same kinds of managerial technologies that allow OpenAI to serve over a hundred million customers with [500 employees](#) will also allow a 500-person political elite, or even a 5-person board, to maintain an iron fist over an entire country. With modern surveillance to collect information, and modern AI to interpret it, there may be no place to hide.

It gets worse when we think about the consequences of AI in warfare. Quoting a [semi-famous post](#) on the philosophy of AI and crypto by [OxAlpha](#):

When there is no need for political-ideological work and war mobilization, the supreme commander of war only needs to consider the situation itself as if it were a game of chess and completely ignore the thoughts and emotions of the pawns/knights/rooks on the chessboard. War becomes a purely technological game.

Furthermore, political-ideological work and war mobilization require a justification for anyone to wage war. Don't underestimate the importance of such "justification". It has been a legitimacy constraint on the wars in human society for thousands of years. Anyone who wants to wage war has to have a reason, or at least a superficially justifiable excuse. You might argue that this constraint is so weak because, in many instances, this has been nothing more than an excuse. For example, some (if not all) of the Crusades were really to occupy land and rob wealth, but they had to be done in the name of God, even if the city being robbed was God's Constantinople. However, even a weak constraint is still a constraint! This little excuse requirement alone actually prevents the warmakers from being completely unscrupulous in achieving their goals. Even an evil like Hitler could not just launch a war right off the bat—he had to spend years first trying to convince the German nation to fight for the living space for the noble Aryan race.

Today, the "human in the loop" serves as an important check on a dictator's power to start wars, or to oppress its citizens internally. Humans in the loop have [prevented](#) nuclear [wars](#), [allowed the opening of the Berlin wall](#), and saved lives during [atrocities](#) like [the Holocaust](#). If armies are robots, this check disappears completely. A dictator could get drunk at 10 PM, get angry at people being mean to them on twitter at 11 PM, and a robotic invasion fleet could cross the border to rain hellfire on a neighboring nation's civilians and infrastructure before midnight.

And unlike previous eras, where there is always some distant corner, where the sky is high and the emperor is far away, where opponents of a regime could regroup and hide and eventually find a way to make things better, with 21st century AI a totalitarian regime may well maintain enough surveillance and control over the world to remain "locked in" forever.

## **d/acc: Defensive (or decentralization, or differential) acceleration**

Over the last few months, the "e/acc" ("effective accelerationist") movement has gained a lot of steam. [Summarized by "Beff Jezos" here](#), e/acc is fundamentally about an appreciation of the truly massive benefits of technological progress, and a desire to accelerate

this trend to bring those benefits sooner.

I find myself sympathetic to the e/acc perspective in a lot of contexts. There's a lot of evidence that [the FDA is far too conservative](#) in its willingness to delay or block the approval of drugs, and bioethics in general far too often seems to operate by the principle that "20 people dead in a medical experiment gone wrong is a tragedy, but 200000 people dead from life-saving treatments being delayed is a statistic". The delays to approving [covid tests and vaccines](#), and [malaria vaccines](#), seem to further confirm this. However, it is possible to take this perspective too far.

In addition to my AI-related concerns, I feel particularly ambivalent about the [e/acc enthusiasm](#) for [military technology](#). In the current context in 2023, where this technology is being made by the United States and immediately applied to defend Ukraine, it is easy to see how it can be a force for good. Taking a broader view, however, enthusiasm about modern military technology as a force for good seems to require believing that the dominant technological power will reliably be one of the good guys in most conflicts, now and in the future

: military technology is good because military technology is being built and controlled by America and America is good. Does being an e/acc require being an America maximalist, betting everything on both the government's present and future morals and the country's future success?

On the other hand, I see the need for new approaches in thinking of how to reduce these risks. The [OpenAI governance structure](#) is a good example: it seems like a well-intentioned effort to balance the need to make a profit to satisfy investors who provide the initial capital with the desire to have a check-and-balance to push against moves that risk OpenAI blowing up the world. In practice, however, their recent [attempt to fire Sam Altman](#) makes the structure seem like an abject failure: it centralized power in an undemocratic and unaccountable board of five people, who made key decisions based on secret information and refused to give any details on their reasoning [until employees threatened to quit](#) en-masse. Somehow, the non-profit board played their hands so poorly that the company's employees [created](#) an [impromptu de-facto union](#)... to side

with the billionaire CEO against them.

Across the board, I see far too many plans to save the world that involve giving a small group of people extreme and opaque power and hoping that they use it wisely. And so I find myself drawn to a different philosophy, one that has detailed ideas for how to deal with risks, but which seeks to create and maintain a more democratic world and tries to avoid centralization as the go-to solution to our problems. This philosophy also goes quite a bit broader than AI, and I would argue that it applies well even in worlds where AI risk concerns turn out to be largely unfounded

. I will refer to this philosophy by the name of d/acc

.

The "d" here can stand for many things; particularly, defense

, decentralization

, democracy

and differential

. First, think of it about defense, and then we can see how this ties into the other interpretations.

## **Defense-favoring worlds help healthy and democratic governance thrive**

One frame to think about the macro consequences of technology is to look at the balance of defense vs offense

. Some technologies make it easier to attack others, in the broad sense of the term: do things that go against their interests, that they feel the need to react to. Others make it easier to defend, and even defend without reliance on large centralized actors.

A defense-favoring world is a better world, for many reasons. First of course is the direct benefit of safety: fewer people die, less economic value gets destroyed, less time is wasted on conflict. What is less appreciated though is that a defense-favoring world makes it easier for healthier, more open and more freedom-respecting forms of governance to thrive.

An obvious example of this is Switzerland

. Switzerland is often considered to be the closest thing the real world has to a classical-liberal governance utopia. Huge amounts of power are devolved to provinces (called "cantons"), major decisions are [decided by referendums](#), and many locals [do not even know who the president is](#). How can a country like this survive [extremely challenging](#) political [pressures](#)? Part of the [answer](#) is [excellent political strategy](#), but the other major part is [very defense-favoring geography](#) in the form of its mountainous terrain.

The flag is a big plus. But so are the mountains.

Anarchist societies in Zomia, famously profiled in James C Scott's new book ["The Art of Not Being Governed"](#), are another example: they too maintain their freedom and independence in large part thanks to mountainous terrain. Meanwhile, the Eurasian steppes are [the exact opposite of a governance utopia](#). Sarah Paine's exposition of [maritime versus continental powers](#) makes similar points, though focusing on water as a defensive barrier rather than mountains. In fact, the combination of ease of voluntary trade and difficulty of involuntary invasion, common to both Switzerland and the island states, seems ideal for human flourishing.

I discovered a related phenomenon when advising quadratic funding experiments within the Ethereum ecosystem: specifically the [Bitcoin Grants funding rounds](#). In [round 4](#), a mini-scandal arose when some of the highest-earning recipients were Twitter influencers, whose contributions are viewed by some as positive and others as negative. My own interpretation of this phenomenon was that there is an imbalance: [quadratic funding](#) allows you to signal that you think something is a public good

, but it gives no way to signal that something is a public bad

. In the extreme, a fully neutral quadratic funding system would fund both sides of a war. And so [for round 5](#), I proposed that Bitcoin should include negative contributions

: you pay \$1 to reduce

the amount of money that a given project receives (and implicitly redistribute it to all other projects). The result [lots](#) of [people](#) hated [it](#)

.

One of the many internet memes that floated around after round 5.

This seemed to me to be a microcosm of a bigger pattern: creating decentralized governance mechanisms to deal with

negative

externalities is socially a very hard problem

. There is a reason why the go-to example of decentralized governance going wrong is mob justice. There is [something about human psychology](#) that makes responding to negatives much more tricky, and much more likely to go very wrong, than responding to positives. And this is a reason why even in otherwise highly democratic organizations, decisions of how to respond to negatives are often left to a centralized board.

In many cases, this conundrum is one of the deep reasons why the concept of "freedom" is so valuable. If someone says something that offends you, or has a lifestyle that you consider disgusting, the pain and disgust that you feel is real, and you may even find it less bad to be physically punched than to be exposed to such things. But trying to agree on what kinds of offense and disgust are socially actionable can have far more costs and dangers than simply reminding ourselves that certain kinds of weirdos and jerks are the price we pay for living in a free society.

At other times, however, the "grin and bear it" approach is unrealistic. And in such cases, another answer that is sometimes worth looking toward is defensive technology. The more that the internet is secure, the less we need to violate people's privacy and use shady international diplomatic tactics to go after each individual hacker. The more that we can build [personalized tools for blocking people on Twitter](#), [in-browser tools for detecting scams](#) and [collective tools for telling apart misinformation](#) and [truth](#), the less we have to fight over censorship. The faster we can make vaccines, the less we have to go after people for being superspreaders. Such solutions don't work in all domains - we certainly don't want a world where everyone has to wear literal body armor - but in domains where we can

build technology to make the world more defense-favoring, there is enormous value in doing so.

This core idea, that some technologies are defense-favoring and are worth promoting, while other technologies are offense-favoring and should be discouraged, has roots in effective altruist literature under a different name: differential technology development

. There is a good [exposition of this principle from University of Oxford researchers](#) from 2022:

Figure 1: Mechanisms by which differential technology development can reduce negative societal impacts.

There are inevitably going to be imperfections in classifying technologies as offensive, defensive or neutral. Like with "freedom", where one can debate whether social-democratic government policies decrease freedom by levying heavy taxes and coercing employers or increase freedom by reducing average people's need to worry about many kinds of risks, with "defense" too there are some technologies that could fall on both sides of the spectrum. Nuclear weapons are offense-favoring, but nuclear power

is [human-flourishing-favoring](#) and offense-defense-neutral. Different technologies may play different roles at different time horizons. But much like with "freedom" (or "equality", or "rule of law"), ambiguity at the edges is not so much an argument against the principle, as it is an opportunity to better understand its nuances.

Now, let's see how to apply this principle to a more comprehensive worldview. We can think of defensive technology, like [other technology](#), as being split into two spheres: the world of atoms

and the world of bits

. The world of atoms, in turn, can be split into micro

(ie. biology, later nanotech) and macro

(ie. what we conventionally think of "defense", but also resilient physical infrastructure). The world of bits I will split on a different axis: how hard is it to agree, in principle, who the attacker is?

. Sometimes it's easy; I call this cyber defense

. At other times it's harder; I call this info defense

.

## Macro physical defense

The most underrated defensive technology in the macro sphere is not even [iron domes](#) (including [Ukraine's new system](#)) and other anti-tech and anti-missile military hardware, but rather resilient physical infrastructure

. The majority of deaths from a nuclear war are likely to come from [supply chain disruptions](#), rather than the initial radiation and blast, and low-infrastructure internet solutions like Starlink have been crucial in [maintaining Ukraine's connectivity](#) for the last year and a half.

Building tools to help people survive and even live comfortable lives independently or semi-independently of long international supply chains seems like a valuable defensive technology, and one with a low risk of turning out to be useful for



offense.

The quest to [make humanity a multi-planetary civilization](#) can also be viewed from a d/acc perspective: having at least a few of us live self-sufficiently on other planets can increase our resilience against something terrible happening on Earth. Even if the full vision proves unviable for the time being, the forms of self-sufficient living

that will need to be developed to make such a project possible may well also be turned to help improve our civilizational resilience on Earth.

## Micro physical defense (aka bio)

Especially due to its [long-term health effects](#), Covid continues to be a [concern](#). But Covid is far from the last pandemic that we will face; there are many aspects of the modern world that make it likely that more pandemics are soon to come:

- Higher population density

makes it much easier for airborne viruses and other pathogens to spread. Epidemic diseases are relatively new in human history and most began with urbanization [only a few thousand years ago](#). [Ongoing rapid urbanization](#) means that population densities will increase further over the next half century.

- Increased air travel

means that airborne pathogens spread very quickly worldwide. People rapidly becoming wealthier means that air travel will likely [increase much

further](<https://www.statista.com/statistics/1089782/forecasted-number-air-passengers-worldwide-generation/>) over the next half century; complexity modeling suggests that [even small increases](#) may have drastic effects. Climate change may increase this risk even further.

- Animal domestication and factory farming

are major risk factors. [Measles](#) probably evolved from a cow virus less than 3000 years ago. [Today's factory farms](#) are also farming new strains of influenza (as well as [fueling antibiotic resistance](#), with consequences for [human innate immunity](#)).

- Modern bio-engineering

makes it easier to create new and more virulent pathogens. Covid [may or may not have leaked](#) from a lab doing intentional "gain of function" research. Regardless, [lab leaks happen all the time](#) and tools are rapidly improving to make it easier to intentionally create extremely deadly viruses, or even [prions \(zombie proteins\)](#). Artificial plagues are particularly concerning in part because [unlike nukes](#), they are unattributable: you can release a virus without anyone being able to tell who created it. It is possible right now

to design a genetic sequence and send it to [awet lab](#) for synthesis, and have it shipped to you within five days.

This is an area where [CryptoRelief](#) and [Balvi](#), two orgs spun up and funded as a result of a large accidental [windfall of Shiba Inu coins](#) in 2021, have been very active. CryptoRelief initially focused on responding to the immediate crisis and more recently has been building up a long-term medical research ecosystem in India, while Balvi has been focusing on moonshot projects to improve our ability to detect, prevent and treat Covid and other airborne diseases. ++Balvi has insisted that projects it funds must be open source++. Taking inspiration from [the 19th century water engineering movement that defeated cholera](#) and other waterborne pathogens, it has funded projects across the whole spectrum of technologies that can make the world more hardened against airborne pathogens by default (see: [update 1](#) and [update 2](#)), including:

- [Far-UVC irradiation R&D](#)
- Air filtering and quality monitoring in [India](#), Sri Lanka, [the United States](#) and elsewhere, and air quality monitoring
- Equipment for cheap and effective [decentralized air quality testing](#)
- Research on Long Covid causes and potential treatment options (the primary cause may be [straightforward](#) but [clarifying mechanisms](#) and finding treatment is harder)
- Vaccines (eg. [RaDVaC](#), [PopVax](#)) and vaccine injury research
- A set of entirely novel non-invasive medical tools
- Early detection of epidemics using analysis of open-source data (eg. [EPIWATCH](#))
- Testing, including very cheap molecular rapid tests
- Biosafety-appropriate masks for when other approaches fail

Other promising areas of interest include [wastewater surveillance of pathogens](#), [improving filtering and ventilation in](#)

[buildings](#), and better understanding and mitigating [risks from poor air quality](#).

There is an opportunity to [build a world](#) that is much more hardened against airborne pandemics, both natural and artificial, by default. This world would feature a highly optimized pipeline where we can go from a pandemic starting, to being automatically detected, to people around the world having access to targeted, [locally-manufacturable and verifiable open source vaccines](#) or [other prophylactics](#), administered via [nebulization](#) or [nose spray](#) (meaning: self-administerable if needed, and no needles required), all within a month. In the meantime, much better air quality would drastically reduce the rate of spread, and prevent many pandemics from getting off the ground at all.

Imagine a future that doesn't have to resort to the sledgehammer of social compulsion - no mandates and worse, and no risk of [poorly designed and implemented mandates that arguably make things worse](#) - because the infrastructure of public health is woven into the fabric of civilization. These worlds are possible, and a medium amount of funding into bio-defense could make it happen. The work would happen even more smoothly if developments are open source, free to users and protected as public goods.

## Cyber defense, blockchains and cryptography

It is generally understood among security professionals that the current state of computer security is pretty terrible. That said, it's easy to understate the amount of progress that has been made. Hundreds of billions of dollars of cryptocurrency are available to anonymously steal by anyone who can hack into users' wallets, and while [far more gets lost or stolen](#) than I would like, it's also a fact that most of it has remained un-stolen for over a decade. Recently, there have been improvements:

- [Trusted hardware chips inside of users' phones

](<https://www.androidauthority.com/titan-m2-google-3261547/>), effectively creating a much smaller high-security operating system inside the phone that can remain protected even if the rest of the phone gets hacked. Among many other use cases, these chips are increasingly being explored as a way to make [more secure crypto wallets](#).

- Browsers as the de-facto operating system

. Over the last ten years, there has been a quiet shift from downloadable applications to in-browser

applications. This has been largely enabled by [WebAssembly \(WASM\)](#). Even Adobe Photoshop, long cited as a major reason why many people cannot practically use Linux because of its necessity and Linux-incompatibility, is now Linux-friendly thanks to being inside the browser. This is also a large security boon: while [browsers do have flaws](#), in general they come with much more sandboxing than installed applications: apps cannot access arbitrary files on your computer.

- Hardened operating systems

. [GrapheneOS](#) for mobile exists, and is very usable. [QubesOS](#) for desktop exists; it is currently somewhat less usable than Graphene, at least in my experience, but it is improving.

- Attempts at moving beyond passwords

. Passwords are, unfortunately, difficult to secure both because they are hard to remember, and because [they are easy to eavesdrop on](#). Recently, there has been a growing movement toward reducing emphasis on passwords, and making multi-factor hardware-based authentication [actually work](#).

However, the lack

of cyber defense in other spheres has also led to major setbacks. The need to protect against spam has led to email becoming [very oligopolistic in practice](#), making it very hard to self-host or create a new email provider. Many online apps, [including Twitter](#), are requiring users to be logged in to access content, and blocking IPs from VPNs, making it harder to access the internet in a way that protects privacy. Software centralization is also risky because of ["weaponized interdependence"](#): the tendency of modern technology to route through centralized choke points, and for the operators of those choke points to use that power to gather information, manipulate outcomes or exclude specific actors - a strategy that seems to even be currently employed [against the blockchain industry itself](#).

These are concerning trends, because it threatens what has historically been one of my big hopes for why the future of freedom and privacy, despite deep tradeoffs, might still turn out to be bright. In his book ["Future Imperfect"](#), David Friedman predicts that we might get a compromise future: the in-person world would be more and more surveilled, but through cryptography, the online world would retain, and even improve, its privacy. Unfortunately, as we have seen, such a counter-trend is far from guaranteed.

This is where my own emphasis on cryptographic technologies such as blockchains and [zero-knowledge proofs](#) comes in

. Blockchains let us create economic and social structures with a "shared hard drive" without having to depend on centralized actors. Cryptocurrency allows individuals to save money and make financial transactions, as they could before the internet with cash, without dependence on trusted third parties that could change their rules on a whim. They can also serve as a fallback anti-sybil mechanism, [making attacks and spam expensive](#) even for users who do not have or do not want to reveal their meat-space identity. Account abstraction, and notably [social recovery wallets](#), can protect our crypto-

assets, and potentially other assets in the future, without over-relying on centralized intermediaries.

Zero knowledge proofs can be used [for privacy](#), allowing users to prove things about themselves without revealing private information. For example, wrap a [digital passport signature](#) in a ZK-SNARK to prove that you are a unique citizen of a given country, without revealing which citizen you are

. Technologies like this can let us maintain the benefits of privacy and anonymity - properties that are widely agreed as being [necessary for applications like voting](#) - while still getting security guarantees and fighting spam and bad actors.

A proposed design for a ZK social media system, where moderation actions can happen and users can be penalized, all without needing to know anyone's identity.

[Zupass](#), incubated at [Zuzalu](#) earlier this year, is an excellent example of this in practice. This is an application, which has already been used by hundreds of people at Zuzalu and more recently by thousands of people for ticketing at [Devconnect](#), that allows you to hold tickets, memberships, ([non-transferable](#)) digital collectibles, and other attestations, and prove things about them all without compromising your privacy. For example, you can prove that you are a unique registered resident of Zuzalu, or a Devconnect ticket holder, without revealing anything else about who you are

. These proofs can be shown in-person, via a QR code, or digitally, to log in to applications like [Zupoll](#), an anonymized voting system available only to Zuzalu residents.

These technologies are an excellent example of d/acc principles: they allow users and communities to verify trustworthiness without compromising privacy, and protect their security without relying on centralized choke points that impose their own definitions of who is good and bad.

They improve global accessibility by creating better and fairer ways to protect a user or service's security than common techniques used today, such as discriminating against entire countries that are deemed untrustworthy. These are very powerful primitives that could be necessary if we want to preserve a decentralized vision of information security going into the 21st century. Working on defensive technologies for cyberspace more broadly can make the internet more open, safe and free in very important ways going forward.

## Info-defense

Cyber-defense, as I have described it, is about situations where it's easy for reasonable human beings to all come to consensus on who the attacker is. If someone tries to hack into your wallet, it's easy to agree that the hacker is the bad guy. If someone tries to DoS attack a website, it's easy to agree that they're being malicious, and are not morally the same as a regular user trying to read what's on the site. There are other situations where the lines are more blurry. It is the tools for improving our defense in these situations that I call "info-defense".

Take, for example, fact checking (aka, preventing "misinformation"). I am a [huge fan of Community Notes](#), which has done a lot to help users identify truths and falsehoods in what other users are tweeting. Community Notes uses a new algorithm which surfaces not the notes that are the most popular

, but rather the notes that are most approved by users across the political spectrum

.

Community Notes in action.

I am also a fan of prediction markets, which can help identify the significance of events in real time

, before the dust settles and there is consensus on which direction is which. The [Polymarket on Sam Altman](#) is very helpful in giving a useful summary of the ultimate consequences of hour-by-hour revelations and negotiations, giving much-needed context to people who only see the individual news items and don't understand the significance of each one.

Prediction markets are often flawed. But Twitter influencers who are willing to confidently express what they think "will" happen over the next year are often even more flawed. There is still room to improve prediction markets much further. For example, a major practical flaw of prediction markets is their low volume on all but the most high-profile events; a natural direction to try to solve this would be to have prediction markets that are played by AIs.

Within the blockchain space, there is a particular type of info defense that I think we need much more of. Namely, wallets should be much more opinionated and active in helping users determine the meaning of things that they are signing, and protecting them from fraud and scams. This is an intermediate case: what is and is not a scam is less subjective than perspectives on controversial social events, but it's more subjective than telling apart legitimate users from DoS attackers or hackers. Metamask has a scam database already, and automatically blocks users from visiting scam sites:

Applications like [Fire](#) are an example of one way to go much further. However, security software like this should not be something that requires explicit installs; it should be part of crypto wallets, or even browsers, by default.

Because of its more subjective nature, info-defense is inherently more collective than cyber-defense: you need to somehow plug into a large and sophisticated group of people to identify what might be true or false, and what kind of application is a

deceptive ponzi. There is an opportunity for developers to go much further in developing effective info-defense, and in hardening existing forms of info-defense. Something like Community Notes could be included in browsers, and cover not just social media platforms but also the whole internet.

## Social technology beyond the "defense" framing

To some degree, I can be justifiably accused of shoehorning by describing some of these info technologies as being about "defense". After all, defense is about helping well-meaning actors be protected from badly-intentioned actors (or, in some cases, from nature). Some of these social technologies, however, are about helping well-intentioned actors form consensus

A good example of this is [pol.is](#), which uses an algorithm similar to Community Notes (and which predates Community Notes) to help communities identify points of agreement between sub-tribes who otherwise disagree on a lot. [Viewpoints.xyz](#) was inspired by pol.is, and has a similar spirit:

Technologies like this could be used to enable more decentralized governance over contentious decisions. Again, blockchain communities are a good testing ground for this, and one where such algorithms have already shown valuable. Generally, decisions over which improvements ("EIPs") to make to the Ethereum protocol are made by a fairly small group in meetings called "[All Core Devs calls](#)". For highly technical decisions, where most community members have no strong feelings, this works reasonably well. For more consequential decisions, which affect protocol economics, or more fundamental values like immutability and censorship resistance, this is often not enough. Back in 2016-17, when a series of contentious decisions around implementing the [DAO fork](#), reducing issuance and (not) [unfreezing the Parity wallet](#), tools like Carbonvote, as well as social media voting, helped the community and the developers to see which way the bulk of the community opinion was facing.

Carbonvote on the DAO fork.

Carbonvote [had its flaws](#): it relied on ETH holdings to determine who was a member of the Ethereum community, making the outcome dominated by a few wealthy ETH holders ("whales"). With modern tools, however, we could make a much better Carbonvote, leveraging multiple signals such as [POAPs](#), Zupass stamps, [Gitcoin passports](#), [Protocol Guild memberships](#), as well as ETH (or even solo-staked-ETH) holdings to gauge community membership.

Tools like this could be used by any community to make higher-quality decisions, find points of commonality, coordinate (physical or digital) migrations or do a number of other things without relying on opaque centralized leadership. This is not defense acceleration per se, but it can certainly be called democracy acceleration. Such tools could even be used to improve and democratize the governance of key actors and institutions working in AI.

## So what are the paths forward for superintelligence?

The above is all well and good, and could make the world a much more harmonious, safer and freer place for the next century. However, it does not yet address the big elephant in the room: superintelligent AI.

The default path forward suggested by many of those who worry about AI essentially leads to a minimal AI world government

. Near-term versions of this include a proposal for a "[multinational AGI consortium](#)" ("MAGIC"). Such a consortium, if it gets established and succeeds at its goals of creating superintelligent AI, would have a natural path to becoming a de-facto minimal world government. Longer-term, there are ideas like the "[pivotal act](#)" theory: we create an AI that performs a single one-time act

which rearranges the world into a game where from that point forward humans are still in charge, but where the game board is somehow more defense-favoring and more fit for human flourishing.

The main practical issue that I see with this so far is that people don't seem to actually trust any specific governance mechanism with the power to build such a thing

. This fact becomes stark when you look at the results to my recent Twitter polls, asking if people would prefer to see AI monopolized by a single entity with a decade head-start, or AI delayed by a decade for everyone:

The size of each poll is small, but the polls make up for it in the uniformity of their result across a wide diversity of sources and options. In nine out of nine cases, the majority of people would rather see highly advanced AI delayed by a decade outright than be monopolized by a single group, whether it's a corporation, government or multinational body

. In seven out of nine cases, delay won by at least two to one. This seems like an important fact to understand for anyone pursuing AI regulation. Current approaches have been focusing on creating licensing schemes and regulatory requirements, trying to restrict AI development to a smaller number of people, but these have seen popular pushback precisely because people don't want to see anyone monopolize something so powerful. Even if such top-down regulatory proposals reduce risks of extinction, they risk increasing the chance of some kind of permanent lock-in to centralized totalitarianism. Paradoxically, could agreements banning extremely advanced AI research outright

(perhaps with exceptions for biomedical AI), combined with measures like mandating

open source for those models that are not banned as a way of reducing profit motives while further improving equality of access, be more

popular?

The main approach preferred by opponents of the "let's get one global org to do AI and make its governance really really good" route is [polytheistic AI](#): intentionally try to make sure there's lots of people and companies developing lots of AIs, so that none of them grows far more powerful than the other

. This way, the theory goes, even as AIs become superintelligent, we can retain a balance of power.

This philosophy is interesting, but my experience trying to ensure "polytheism" within the Ethereum ecosystem does make me worry that this is an inherently unstable equilibrium. In Ethereum, we have intentionally tried to ensure decentralization of many parts of the stack: ensuring that there's no single codebase that controls [more than half of the proof of stake network](#), trying to counteract the [dominance of large staking pools](#), improving [geographic decentralization](#), and so on. Essentially, Ethereum is actually attempting to execute on the old libertarian dream of a market-based society that uses social pressure, rather than government, as the antitrust regulator. To some extent, this has worked: the [Prism client's dominance](#) has dropped from above 70% to under 45%. But this is not some automatic market process: it's the result of human intention and coordinated action.

My experience within Ethereum is mirrored by learnings from the broader world as a whole, where many markets have proven to be [natural monopolies](#). With superintelligent AIs acting independently of humans, the situation is even more unstable. Thanks to [recursive self-improvement](#), the strongest AI may pull ahead very quickly, and once AIs are more powerful than humans, there is no force that can push things back into balance.

Additionally, even if we do

get a polytheistic world of superintelligent AIs that ends up stable, we still have the other

problem: that we get a universe where humans are pets.

## A happy path: merge with the AIs?

A different option that I have heard about more recently is to focus less on AI as something separate from humans, and more on tools that enhance

human cognition rather than replacing

it

.

One near-term example of something that goes in this direction is AI drawing tools. Today, the most prominent tools for making AI-generated images only have one step at which the human gives their input, and AI fully takes over from there. An alternative would be to focus more on AI versions of Photoshop: tools where the artist or the AI might make an early draft of a picture, and then the two collaborate on improving it with a process of real-time feedback.

Photoshop generative AI fill, 2023. [Source](#). I tried, it and it takes time to get used to but it actually works quite well!

Another direction in a similar spirit is the [Open Agency Architecture](#), which proposes splitting the different parts of an AI "mind" (eg. making plans, executing on plans, interpreting information from the outside world) into separate components, and introducing diverse human feedback in between these parts.

So far, this sounds mundane, and something that almost everyone can agree that it would be good to have. The economist Daron Acemoglu's work is far from this kind of AI futurism, but his new book [Power and Progress](#) hints at wanting to see more of exactly these types of AI.

But if we want to extrapolate this idea of human-AI cooperation further, we get to more radical conclusions

. Unless we create a world government powerful enough to detect and stop every small group of people hacking on individual GPUs with laptops, someone is going to create a superintelligent AI eventually - one that can think a [thousand times faster](#) than we can - and no combination of humans using tools with their hands is going to be able to hold its own against that. And so we need to take this idea of human-computer cooperation much deeper and further.

A first natural step is [brain-computer interfaces](#)

. Brain-computer interfaces can give humans much more direct access to more-and-more powerful forms of computation and cognition, reducing the two-way communication loop between man and machine from seconds to milliseconds. This would also greatly reduce the "mental effort" cost to getting a computer to help you gather facts, give suggestions or execute



on a plan.

Later stages of such a roadmap admittedly get weird. In addition to brain-computer interfaces, there are various paths to improving our brains directly through innovations in biology. An eventual further step, which merges both paths, may involve [uploading our minds](#) to run on computers directly. This would also be the ultimate d/acc for physical security: protecting ourselves from harm would no longer be a challenging problem of protecting inevitably-squishy human bodies, but rather a much simpler problem of making data backups.

Directions like this are sometimes met with worry, in part because they are irreversible, and in part because they may give powerful people more advantages over the rest of us. Brain-computer interfaces in particular have dangers - after all, we are talking about literally reading and writing to people's minds

. These concerns are exactly why I think it would be ideal for a leading role in this path to be held by a security-focused open-source movement, rather than closed and proprietary corporations and venture capital funds. Additionally, all of these issues are worse

with superintelligent AIs that operate independently from humans, than they are with augmentations that are closely tied to humans. The divide between "enhanced" and "unenhanced" already exists today due to [limitations in who can and can't use ChatGPT](#).

If we want a future that is both superintelligent and

"human", one where human beings are not just pets, but actually retain meaningful agency over the world, then it feels like something like this is the most natural option

. There are also good arguments why this could be a safer AI alignment path: by involving human feedback at each step of decision-making, we reduce the incentive to offload high-level planning responsibility to the AI itself, and thereby reduce the chance that the AI does something totally unaligned with humanity's values on its own.

One other argument in favor of this direction is that it may be more socially

palatable than simply shouting "[pause AI](#)" without a complementary message providing an alternative path forward. It will require a philosophical shift from the current mentality that tech advancements that touch humans are dangerous but advancements that are separate from humans are by-default safe. But it has a huge countervailing benefit: it gives developers something to do

. Today, the AI safety movement's primary message to AI developers seems to be [you should just stop](#)". One can [work on alignment research](#), but today this lacks economic incentives. Compared to this, the common e/acc message of "you're already a hero just the way you are" is understandably extremely appealing. A d/acc message, one that says "you should build, and build profitable things, but be much more selective and intentional in making sure you are building things that help you and humanity thrive", may be a winner.

## Is d/acc compatible with your existing philosophy?

- If you are an e/acc

, then d/acc is a subspecies of e/acc - just one that is much more selective and intentional.

- If you are an [effective altruist

](<https://www.effectivealtruism.org/>), then d/acc is a re-branding of the effective-altruist idea of [differential technology development](#), though with a greater emphasis on liberal and democratic values.

- If you are a libertarian

, then d/acc is a sub-species of techno-libertarianism, though a more pragmatic one that is more critical of "the techno-capital machine", and willing to accept government interventions today (at least, if cultural interventions don't work) to prevent much worse un-freedom tomorrow.

- If you are a Pluralist

, in the [Glen Weyl sense of the term](#), then d/acc is a frame that can easily include the emphasis on better democratic coordination technology that Plurality values.

- If you are a public health advocate

, then d/acc ideas can be a source of a broader long-term vision, and opportunity to find common ground with "tech people" that you might otherwise feel at odds with.

- If you are a blockchain advocate

, then d/acc is a more modern and broader narrative to embrace than the fifteen-year-old emphasis on hyperinflation and banks, which puts blockchains into context as one of many tools in a concrete strategy to build toward a brighter future.

- If you are a [solarpunk](#)

, then d/acc is a subspecies of solarpunk, and incorporates a similar emphasis on intentionality and collective action.

- If you are a [lunarpunk](#)

, then you will appreciate the d/acc emphasis on informational defense, through maintaining privacy and freedom.

## We are the brightest star

I love technology because technology expands human potential. Ten thousand years ago, we could build some hand tools, change which plants grow on a small patch of land, and [build basic houses](#). Today, we can build [800-meter-tall towers](#), store the entirety of recorded human knowledge in a device we can hold in our hands, communicate instantly across the globe, double our lifespan, and live happy and fulfilling lives without fear of our best friends regularly dropping dead of disease.

We started from the bottom, now we're here.

I believe that these things are deeply good, and that expanding humanity's reach even further to the planets and stars is deeply good, because I believe humanity is deeply good

. It is fashionable in some circles to be skeptical of this: the [voluntary human extinction movement](#) argues that the Earth would be better off without humans existing at all, and many more want to see [much smaller number of human beings](#) see the light of this world in the centuries to come. It is common to [argue](#) that [humans are bad](#) because we cheat and steal, engage in colonialism and war, and mistreat and annihilate other species. My reply to this style of thinking is one simple question: compared to what?

Yes, human beings are often mean, but we much more often show kindness and mercy, and work together for our common benefit. Even during wars we often take care to protect civilians - certainly not nearly enough, but also far more than [we did 2000 years ago](#). The next century may well bring widely available non-animal-based meat, eliminating the [largest moral catastrophe](#) that human beings can justly be blamed for today. Non-human animals are not like this. There is no situation where a cat will adopt an entire lifestyle of refusing to eat mice as a matter of ethical principle. The Sun is growing brighter every year, and in about [one billion years](#), it is expected that this will make the Earth too hot to sustain life. Does the Sun even think

about the genocide that it is going to cause?

And so it is my firm belief that, out of all the things that we have known and seen in our universe, we, humans, are the brightest star

. We are the one thing that we know about that, even if imperfectly, sometimes make an earnest effort to care about "the good", and adjust our behavior to better serve it. Two billion years from now, if the Earth or any part of the universe still bears the beauty of Earthly life, it will be human artifices like space travel and [geoengineering](#) that will have made it happen.

We need to build, and accelerate. But there is a very real question that needs to be asked: what is the thing that we are accelerating towards? The 21st century may well be [the

pivotal century](<https://www.cold-takes.com/most-important-century/>) for humanity, the century in which our fate for millennia to come gets decided. Do we fall into one of a number of traps from which we cannot escape, or do we find a way toward a future where we retain our freedom and agency? These are challenging problems. But I look forward to watching and participating in our species' grand collective effort to find the answers.