



FROM EXTROPY TO AGI

The Ideas Behind Silicon
Valley's AI Future

Transhumanism · Singularity · Rationalism · Effective Altruism
Longtermism · AI Safety · TESCREAL · e/acc

A documented intellectual history: curious, critical, and
pro-AI without treating speculation as fact.

Research article · October 2026

FROM EXTROPY TO AGI

The Ideas Behind Silicon Valley's AI Future

How cybernetics, counterculture, transhumanism, singularity thinking, rationalism, effective altruism, longtermism, AI safety and accelerationism became part of the strange intellectual weather around modern artificial intelligence.

HOW TO READ THIS ARTICLE

This is an intellectual genealogy, not a conspiracy theory. The people and movements discussed here overlap in real, documented ways, but they do not form one secret organization, one political program or one set of beliefs. When the article describes a criticism - especially the TESCREAL framework - it attributes that criticism rather than presenting it as settled fact. The goal is to understand how ideas travel into institutions, not to declare everyone in Silicon Valley a member of the same club.

Silicon Valley is often described as if it were simply a place where engineers build products and venture capitalists fund them. That picture is incomplete. For decades, the region has also been an unusually dense laboratory for ideas about what human beings are, what technology is for, whether biological limits should be accepted, whether intelligence can be engineered, and who gets to decide what the future should look like.

Some of those ideas are serious science. Some are moral philosophy. Some are political ideology. Some began as science fiction. Some are now corporate strategy. And some are so grand that, read cold, they sound as if a philosophy seminar, a startup pitch deck and a space opera were accidentally merged into one PDF.

The interesting question is therefore not, 'Is Silicon Valley a cult?' It is: how did a collection of fringe and mainstream ideas - often disagreeing with one another - become influential enough that today's frontier AI companies talk openly about AGI, superintelligence, alignment, civilization-scale risk and the long-run future of humanity?

THE SHORT ANSWER

Silicon Valley did not inherit one master ideology. It inherited several streams of thought about tools, freedom, intelligence, human enhancement and the far future. Dense social networks connected them; unusually large pools of capital turned some of them into institutions; and rapid AI progress made old speculative arguments newly consequential.

Inside this article

1. AI, AGI and ASI - the three letters we need before anything else
2. Before Silicon Valley: computers become partners, not just calculators
3. The Bay Area learns to treat technology as liberation
4. Transhumanism: what if the human body is version 1.0?
5. Max More, Extropy and the 1990s futurist laboratory
6. Digital cash, cryonics and the strange usefulness of weird communities
7. The singularity: when forecasting becomes a theory of history
8. Yudkowsky, MIRI and the pivot from "build it" to "align it"
9. Rationalism: a community built around thinking about thinking
10. Effective altruism starts somewhere else - and then the streams meet
11. Longtermism and existential risk: the moral horizon becomes enormous
12. Money, networks and institutions: philosophy gains a budget
13. DeepMind, OpenAI and Anthropic: the ideas enter frontier labs
14. e/acc: the counter-revolt against caution
15. TESCREAL: a useful critical lens, and where it can overreach
16. Is it a cult? Mostly no. Why can it feel cult-shaped?
17. Belief, branding and the marketing of inevitability
18. What is actually happening in Silicon Valley?
19. Timeline, glossary and source notes

START HERE

1. AI, AGI and ASI - the three letters we need before anything else

Artificial intelligence, in the broad everyday sense, means computer systems that perform tasks we associate with intelligence: recognizing patterns, understanding language, generating images, planning, predicting, writing code, playing games or helping people make decisions. Today's AI can be remarkably general in practice, but it is still built from statistical and computational methods rather than from a little human-shaped mind hiding inside the server. ^[34]

AGI means artificial general intelligence. The phrase is important and slippery. There is no universally accepted scientific test for it. A rough everyday meaning is an AI that can learn, reason and apply knowledge across a very wide range of domains at something like human-level competence or beyond. Different institutions use different operational definitions; for example, OpenAI's charter historically framed AGI in economic-performance terms, while researchers elsewhere use capability-based definitions. ^{[34][35]}

ASI means artificial superintelligence: a hypothetical system whose general cognitive abilities substantially exceed those of humans. The distinction matters. AGI is usually the threshold question - when does machine intelligence become broadly comparable to human intelligence? ASI is the next question - what happens if it becomes much more capable than us?

NORMAL-HUMAN TRANSLATION

AI: the broad field and today's systems. AGI: a disputed name for very general, human-level-or-better machine capability. ASI: hypothetical machine intelligence far beyond human capability. None of these terms automatically implies consciousness, feelings, a robot body or evil intentions.

By 2026 these are not merely fan-forum words. OpenAI still defines its mission around AGI; Google DeepMind says AGI could drive one of history's greatest transformations and in June 2026 published a report explicitly examining the path 'From AGI to ASI'; Anthropic describes its mission in terms of navigating the transition through transformative AI safely. That does not prove AGI or ASI is imminent. It does show that language once associated mainly with futurists now lives inside the institutions building frontier systems. ^{[31][32][33][25]}

1950S-1960S

2. Before Silicon Valley: computers become partners, not just calculators

The story begins before personal computers, startups or even the modern Silicon Valley myth. Early computers were largely associated with governments, universities, corporations and military research. They were expensive machines used for calculation, logistics and data processing. Yet a few researchers were already imagining something much more intimate: computers as partners in human thought.

In 1960, psychologist and computer scientist J. C. R. Licklider described 'man-computer symbiosis': a future in which people set goals, formulate hypotheses and make judgments while computers handle routinizable work and help humans think more effectively. Two years later, Douglas Engelbart's landmark report on 'augmenting human intellect' treated the human plus tools, concepts and methods as a system that could be deliberately improved. ^{[1][2]}

This is a crucial ancestor of today's AI conversation. The earliest dream was not necessarily 'replace the human.' It was often 'amplify the human.' The computer was imagined as an intellectual prosthesis - something that expands what people can understand and do.

A RECURRING FORK IN THE ROAD

Much of AI history oscillates between two visions: AUGMENTATION - technology makes humans more capable; and AUTONOMY - technology increasingly performs cognitive work itself. Modern AI products often contain both at once.

1960S-1990S

3. The Bay Area learns to treat technology as liberation

The second ingredient is cultural rather than technical. Historian Fred Turner traces how parts of the American counterculture - especially networks around Stewart Brand and the Whole Earth Catalog - helped recast computers from symbols of centralized bureaucracy into tools of personal empowerment, self-sufficiency, community and decentralization. The Whole Earth Catalog, first published in 1968, famously organized itself around 'access to tools.' ^{[3][4]}

This did not mean that hippies simply invented Silicon Valley. The history is more complicated. But a powerful cultural translation occurred: technology could be anti-bureaucratic. Networks could weaken hierarchy. Personal computers could give individuals capabilities once reserved for institutions. Online communities could become new forms of social organization. ^[3]

By the mid-1990s, critics Richard Barbrook and Andy Cameron gave this mixture a name: 'the Californian Ideology.' Their argument was that Bay Area digital culture blended countercultural anti-authoritarianism with market individualism and technological determinism. In their telling, ideas associated with both the New Left and the libertarian right could meet inside one sunny belief: networked technology would route around old institutions and produce a freer future. ^[5]

Whether one accepts their political critique or not, the descriptive insight remains useful. Silicon Valley did not inherit a neutral relationship with technology. It inherited a habit of treating technology as a vehicle for social transformation.

A family tree, not a conspiracy chart

Documented streams of ideas that overlap, split, and recombine.

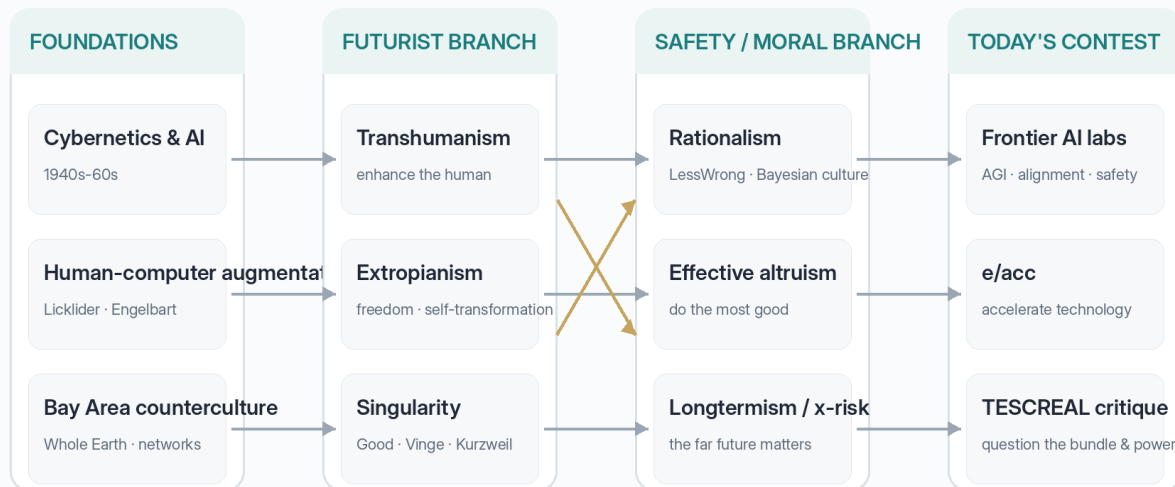


Figure 1. A simplified intellectual genealogy. Real history contains many more cross-links, critics and independent origins.

THE HUMAN BECOMES AN ENGINEERING PROJECT

4. Transhumanism: what if the human body is version 1.0?

Transhumanism is the broad idea that humans can and perhaps should use science and technology to overcome major biological limitations. Depending on the transhumanist, that can include curing disease, extending healthy lifespan, enhancing cognition, changing bodies, integrating with machines, redesigning reproduction or eventually becoming something recognizably posthuman.^[6]

That umbrella is politically and philosophically diverse. Some transhumanists emphasize individual freedom. Some emphasize social access and equality. Some are libertarian; others are social democratic or otherwise politically distant from the early Californian scene. This is important because later critics sometimes talk about 'transhumanism' as if it were one doctrine. It never was.^[6]

The recurring transhumanist move is simple: a limitation that earlier cultures treated as fate becomes a technical problem. Bad eyesight? Correct it. A failing heart? Replace or repair it. Genetic disease? Perhaps edit it. Cognitive limits? Augment them. Aging? Research it as biology rather than destiny. The philosophical disagreement begins when the list moves from healing to redesigning what counts as a human being.

THE ONE-SENTENCE VERSION

Transhumanism asks: if medicine already changes the limits of human biology, where exactly should enhancement stop - and who gets to draw the line?

5. Max More, Extropy and the 1990s futurist laboratory

Modern transhumanism acquired a distinctive West Coast form through Max More and the Extropian community. In 1990 More's Extropian Principles celebrated boundless expansion, self-transformation, dynamic optimism and intelligent technology, while also opposing centralized social control. The early version had an unmistakably libertarian flavor, though More later revised parts of the framework. ^{[7][6]}

The journal Extropy, published from 1988 to 1996, now reads like a time capsule from a future that partly arrived. Across seventeen issues, its contributors discussed life extension, neural computation, nanotechnology, uploading, privacy, digital cash, prediction markets, decentralized governance and posthuman possibilities. Not every idea was correct. The striking part is the breadth: many subjects that later became separate Silicon Valley industries or research fields were being discussed in one small intellectual scene. ^[8]

Bostrom's own history of transhumanism records another revealing fact. When he and David Pearce founded the World Transhumanist Association in 1998, one aim was to create a more mature, academically respectable form of transhumanism and distance it from what some critics had regarded as the 'cultishness' of earlier gatherings. In other words, even insiders understood that the movement's social style could appear unusual from the outside. ^[6]

YES, THIS IS WHERE IT GETS DELIGHTFULLY WEIRD

A 1990s futurist scene could contain serious discussion of cryptography, neurotechnology, free markets, indefinite life extension, mind uploading and space expansion - sometimes before lunch. The mistake is assuming that 'weird' automatically means 'irrelevant.'

IDEAS TRAVEL THROUGH PEOPLE

6. Digital cash, cryonics and the strange usefulness of weird communities

Small intellectual communities matter because they connect people who would otherwise work in different fields. Extropy is a good example. In 1993, Hal Finney published an article in Extropy about protecting privacy with electronic cash. Finney later became one of the most important early figures in Bitcoin history. The point is not that Extropy invented Bitcoin. It did not. The point is that the same networks discussing posthumanism were also incubating serious interest in cryptographic money and privacy. ^[9]

Cryonics was another persistent element of this world. Max More joined Alcor Life Extension Foundation in the 1980s and became its CEO in 2011. Cryonics attempts to preserve legally dead people at very low temperatures in the hope that future technology might one day make repair and revival possible. No cryopreserved human has ever been revived; the idea remains radically speculative. But philosophically it fits the Extropian pattern perfectly: death is treated not as sacred inevitability but as a problem whose technical difficulty may change over time. ^[10]

Here is the useful lesson. Fringe communities can be wrong about many things and still be early about others. Being unconventional increases the chance of entertaining nonsense; it also increases the chance of examining ideas that mainstream institutions dismiss too quickly. The hard part is telling which is which before history answers.

1965-2000S

7. The singularity: when forecasting becomes a theory of history

A separate but increasingly intertwined tradition concerned machine intelligence itself. In the 1960s, mathematician I. J. Good described the possibility of an 'ultraintelligent' machine that could help design better machines, potentially creating an intelligence explosion. The idea is recursive: once the thing improving AI is itself AI, the pace of improvement might change dramatically. ^[11]

Science-fiction author and computer scientist Vernor Vinge gave the modern singularity idea much of its cultural power in 1993. He argued that the creation of superhuman intelligence would produce a transition so profound that the world beyond it would be difficult for present-day humans to predict. Vinge's timetable was far too confident by ordinary forecasting standards, but the conceptual frame survived: advanced machine intelligence might not be just another technology. It might be a break in historical continuity. ^[12]

Ray Kurzweil later popularized the singularity for a mass audience by combining AI, biotechnology and exponential technological progress. At this point the story acquires an almost theological shape while remaining secular: there is a before, a transformative event, and an after in which intelligence, bodies, longevity and society may be radically altered.

WHY THE WORD "SINGULARITY"?

It is borrowed by analogy from domains where our normal predictive framework stops working. In futurism, it means a hypothetical point after which technological change - especially superhuman AI - becomes so fast or profound that ordinary forecasting breaks down.

2000S

8. Yudkowsky, MIRI and the pivot from "build it" to "align it"

Eliezer Yudkowsky's early organization - founded in 2000 as the Singularity Institute for Artificial Intelligence, later renamed the Machine Intelligence Research Institute or MIRI - began with a goal that now surprises many people: accelerate progress toward smarter-than-human AI. MIRI's own history says Yudkowsky initially assumed greater intelligence would bring greater morality. ^[13]

Then came the pivot. Between roughly 2000 and 2003, Yudkowsky became convinced that intelligence and values are separate problems. A system can be brilliant at achieving goals without sharing human goals. The institute therefore moved toward what was then called 'Friendly AI' and

what later became part of the broader alignment problem: how do we make highly capable AI systems reliably pursue intentions compatible with human values and control?^[13]

This was not initially mainstream machine learning. It was closer to theoretical futurism, decision theory and philosophy. But the basic question is now ordinary enough that major AI companies employ large safety, alignment, interpretability and control teams. One of the strangest transformations in this history is watching a concern migrate from obscure mailing lists into corporate research departments.

ALIGNMENT, WITHOUT THE MYSTIQUE

If you ask an AI to optimize something, how do you make sure it reliably does what you actually meant - including in new situations, at high capability, and when shortcuts or unintended strategies are available? That practical problem exists long before any hypothetical superintelligence.

OVERCOMING BIAS → LESSWRONG

9. Rationalism: a community built around thinking about thinking

Yudkowsky's response to the difficulty of AI safety was partly social: if the problem required unusually careful reasoning, perhaps a community should train itself to reason better. The group blog Overcoming Bias began in 2006, and LessWrong followed in 2009. Its culture emphasized Bayesian reasoning, cognitive biases, decision theory, calibration and what participants often called 'the art of rationality.'^[14]

The community did not stay purely online. LessWrong's own histories describe an increasingly physical Bay Area network in which people attended meetups, lived together, became friends, dated, founded organizations and worked on shared problems. That sociological fact matters because ideas travel much faster through dense human networks than through papers alone.^[14]

Rationalism also inherited personnel and themes from earlier transhumanist and singularity communities. Again, this does not mean every rationalist is a transhumanist or an AI doomer. It means the intellectual genealogies overlap in traceable ways.

MORAL PHILOSOPHY ENTERS THE NETWORK

10. Effective altruism starts somewhere else - and then the streams meet

Effective altruism, or EA, deserves separate treatment because it did not begin as a Silicon Valley singularity doctrine. One important philosophical ancestor is Peter Singer's 1972 argument that people with resources have strong moral reasons to prevent serious suffering when they can do so at relatively modest cost. Decades later, organizations such as GiveWell, Giving What We Can and 80,000 Hours pushed a more empirical question: not just 'should we help?' but 'which actions help the most per unit of money, time or career effort?'^{[15][16]}

The early EA emphasis included global health, poverty reduction, evidence-based charity and animal welfare. It was not originally synonymous with AI extinction risk. EA's own history says the movement formed through several communities coming together, including Oxford organizations, GiveWell and the Bay Area rationality community. That convergence is the key historical moment. ^[16]

The same optimization instinct that asks which charity saves the most lives can be applied to much larger questions. If a low-probability catastrophe could end civilization entirely, should preventing it outrank ordinary present-day problems? If the future could contain enormous numbers of people, should their welfare count in today's moral calculations? These questions pulled some parts of EA toward existential risk and the far future.

DO NOT MERGE THESE LABELS TOO QUICKLY

Someone can support evidence-based global-health charity without being a longtermist. A longtermist can disagree with transhumanism. A rationalist can reject EA. A transhumanist can dislike AI safety culture. The overlap is real; the identities are not interchangeable.

OXFORD AND THE FAR FUTURE

11. Longtermism and existential risk: the moral horizon becomes enormous

Nick Bostrom's 2002 paper on existential risk gave a formal name to a category of catastrophe that would not merely kill many people but permanently destroy or drastically curtail humanity's future potential. His 2003 essay 'Astronomical Waste' pushed the utilitarian logic further: if advanced civilization could eventually support vast numbers of valuable lives, then preserving the possibility of that future could have enormous moral importance. ^{[17][18]}

Longtermism later developed this idea more systematically. In one influential formulation, Hilary Greaves and William MacAskill argued that positively affecting the long-run future can be a key moral priority. The argument is not simply 'future people matter.' Most moral views can accept that. The stronger claim is that the scale of the future may make long-run effects dominate many decisions today. ^[19]

This is where the arithmetic becomes both powerful and controversial. Expected-value reasoning can make a tiny probability of an enormous future outcome compete with certain harms happening now. Supporters see this as taking the future seriously. Critics worry that highly speculative numbers can overwhelm concrete present-day evidence and shift influence toward people who claim to know how to protect an unimaginably large future.

A SIMPLE EXAMPLE OF THE CONTROVERSY

Helping 10,000 people today is concrete. Reducing a 0.01% chance of a catastrophe affecting billions - or all future generations - may look larger in expected-value calculations. The mathematics can be straightforward; the difficult part is whether the probabilities and assumptions deserve that confidence.

WHEN IDEAS BECOME INFRASTRUCTURE

12. Money, networks and institutions: philosophy gains a budget

Most philosophies remain books, classrooms and arguments. Silicon Valley philosophies can become grant programs, startups, research institutes and venture portfolios. That difference is enormous.

Good Ventures, founded by Cari Tuna and Dustin Moskovitz, began collaborating with GiveWell in the early 2010s. The project that became Open Philanthropy - renamed Coefficient Giving in 2025 - expanded beyond global health into areas including scientific research, pandemic preparedness and global catastrophic risks. Its own history says that by mid-2025 it had directed more than \$4 billion in total giving.^[20]

Peter Thiel provides another bridge between futurist ideas and technology capital. A well-documented episode from the 2010 Singularity Summit shows DeepMind co-founders Demis Hassabis and Shane Legg meeting Thiel through Yudkowsky's network; Thiel subsequently became the young company's first major outside investor. This does not mean DeepMind was 'created by Extropians.' It demonstrates something more ordinary and more important: conferences and intellectual networks can route capital toward ambitious technical projects.^[21]

In January 2015 Elon Musk committed \$10 million to the Future of Life Institute for research aimed at keeping AI beneficial. The timing matters: this was years before ChatGPT and before today's consumer generative-AI boom. Whatever one thinks of later marketing, serious concern about advanced-AI risk was already attracting real money before frontier AI became a mass-market product.^[22]

Oxford's Future of Humanity Institute became a major hub for work on existential risk and transformative technology before closing in April 2024. Its own final history says that questions once pursued at the margins of a philosophy department are now pursued by leading AI labs, governments, nonprofits and specialized research centers. That may be the clearest summary of this entire intellectual migration.^[30]

FROM PHILOSOPHY TO COMPANY MISSION

13. DeepMind, OpenAI and Anthropic: the ideas enter frontier labs

DeepMind was founded with an explicit ambition to build general-purpose artificial intelligence. The intellectual context around its earliest financing included the Singularity Summit world, but DeepMind also grew from mainstream machine learning, neuroscience and games research. It would be misleading to reduce the company to one philosophical lineage. The important point is that AGI - once a niche futurist phrase - became a serious corporate research target.^{[21][33]}

OpenAI's December 2015 launch post described a nonprofit research company intended to advance digital intelligence in a way that would benefit humanity as a whole. It explicitly acknowledged both the potential benefits and the potential damage of human-level AI. This was mission language, but it was also a statement of worldview: advanced AI was framed from the beginning as a civilization-scale technology.^[23]

Anthropic, founded in 2021 by former OpenAI researchers, was even more explicit about safety. Its first funding announcement called it an AI safety and research company focused on reliable, steerable and interpretable general AI systems. Its current constitution describes the company as occupying a peculiar position: it considers transformative AI potentially dangerous while developing the technology itself, reasoning that safety-focused labs should remain at the frontier. ^{[24][25]}

By 2026, the vocabulary has become remarkably direct. OpenAI writes about helping navigate a transition to a 'post-AGI world' and a long-term goal of personal AGI. Google DeepMind's current mission speaks openly about AGI and its June 2026 research considers pathways from AGI to ASI. These are proposals and forecasts, not proof that either milestone has been reached. But linguistically, the distance between 1990s futurist forums and 2026 corporate strategy has collapsed. ^{[31][32][33]}

THE IRONY

A recurring pattern is: 'This technology may become dangerously powerful, therefore safety-minded people need to help build it.' Critics see a conflict of interest. Supporters see responsible participation in an inevitable race. Both interpretations can coexist because they concern motive, incentives and strategy rather than a simple factual dispute.

2022 ONWARD

14. e/acc: the counter-revolt against caution

Effective accelerationism - usually written e/acc - emerged around 2022 partly as a reaction against the caution, longtermism and existential-risk emphasis associated with parts of EA and AI-safety culture. Its early writings combine technological optimism with a 'physics-first' language of thermodynamics, adaptation, competition and increasing complexity. ^[26]

One of its best-known early advocates used the pseudonym 'Based Beff Jezos,' later identified as physicist and entrepreneur Guillaume Verdon. e/acc's tone is deliberately irreverent: build, accelerate, decentralize, resist centralized gatekeeping, and treat technological progress as a positive evolutionary force. It is less a tidy academic doctrine than a networked cultural movement with manifestos, memes and venture-capital appeal. ^[27]

Marc Andreessen's 2023 Techno-Optimist Manifesto is not identical to e/acc, but it occupies nearby territory. It calls intelligence the ultimate engine of progress, describes human-machine cybernetic systems and predicts an intelligence takeoff. The old themes return: intelligence expands capability; technological progress is morally valuable; stagnation is dangerous; abundance is preferable to restraint. ^[28]

THE FAMILY ARGUMENT

AI-safety culture says: powerful AI may be extraordinary, so control the risk. e/acc says: powerful technology may be extraordinary, so do not let fear or centralized control choke progress. Their fight is fierce partly because both sides already grant technology enormous historical importance.

THE ACRONYM YOU WILL NOW SEE EVERYWHERE

15. TESCREAL: a useful critical lens, and where it can overreach

TESCREAL is an acronym coined by critics to group Transhumanism, Extropianism, Singularitarianism, Cosmism, Rationalism, Effective Altruism and Longtermism. Timnit Gebru and Emile P. Torres argue that this bundle helps explain why parts of the AI industry treat building AGI and managing its cosmic-scale consequences as a natural project. Their 2024 paper also advances a much stronger historical argument linking this worldview to traditions of eugenic thought and to the concentration of power.^[29]

The acronym is valuable because it points to real overlap that is easy to miss when each movement is studied separately. Bostrom's own history confirms direct organizational and social continuity between Extropianism and later transhumanism. LessWrong's histories confirm recruitment from transhumanist communities. EA's own account confirms convergence with the Bay Area rationality community. The links are not imaginary.^{[6][14][16]}

But 'TESCREAL' becomes misleading if it is treated as a neutral scientific category or a single organization. These movements disagree about politics, moral theory, risk, markets, AI development and even what counts as a good future. A malaria-focused effective altruist may have little interest in mind uploading. An e/acc advocate may despise longtermist caution. A transhumanist may favor democratic regulation. A rationalist may reject the whole AGI narrative.

The best use of TESCREAL, then, is as a critical question rather than a stamp: when AGI rhetoric invokes humanity, intelligence, progress, the far future or posthuman possibility, which older ideas are being imported - and whose interests are served by that framing?

EDITORIAL RULE FOR THIS ARTICLE

TESCREAL is treated here as a contested analytical framework. The documented historical overlaps are facts; the claim that they form one coherent ideology - and especially claims about their moral or political meaning - are interpretations that require argument.

THE QUESTION EVERYONE EVENTUALLY ASKS

16. Is it a cult? Mostly no. Why can it feel cult-shaped?

Calling the whole ecosystem a cult is inaccurate. There is no single leader, doctrine, initiation process or central organization. The movements split constantly. Some want AI slowed; others want it accelerated. Some are utilitarians; others are libertarians; others reject both. OpenAI and Anthropic compete commercially while articulating different safety philosophies. That is a very odd cult structure if everyone keeps founding rival institutions.

Still, parts of the ecosystem can feel cult-shaped because high-commitment intellectual communities often develop recognizable social features: canonical texts, specialist vocabulary, conferences, house communities, charismatic thinkers, moral urgency, insider jokes, career networks and stories in which the group is working on something civilization-defining. None of those

characteristics is sufficient to make a cult. Universities, religions, political movements, open-source communities and fandoms can display several of them too.

The historically fair point is that even Bostrom's sympathetic account of transhumanism acknowledged that some critics had regarded earlier gatherings as 'cultish,' and that the later World Transhumanist Association deliberately sought a more academically respectable presentation. The perception is therefore not a twenty-first-century social-media invention.^[6]

BETTER VOCABULARY

Instead of asking only 'Is this a cult?', ask: How socially closed is the community? How does it handle dissent? Are claims falsifiable? Are conflicts of interest disclosed? Can members revise beliefs without losing status? Those questions tell us much more.

WHERE PHILOSOPHY MEETS THE PITCH DECK

17. Belief, branding and the marketing of inevitability

The history makes one easy explanation difficult to sustain: the entire AGI-risk story was not invented after ChatGPT as a marketing campaign. Good wrote about machine intelligence explosion in the 1960s; Vinge wrote about superhuman intelligence in 1993; MIRI pivoted toward AI alignment in the early 2000s; Bostrom formalized existential risk in 2002; Musk funded beneficial-AI research in early 2015; OpenAI launched later that year with an explicitly civilization-scale mission. The beliefs predate the current products.^{[11][12][13][17][22][23]}

But sincerity and marketing are not opposites. A company can sincerely believe a technology is transformative and also benefit enormously from saying so. 'We make useful software' positions a company in a competitive market. 'We are building a technology that could reshape civilization' positions it as a historical actor. The second story attracts capital, talent, policy attention, media coverage and moral significance.

The most powerful rhetorical move may be inevitability. If AGI is framed as something that will be built somewhere no matter what, the public question subtly changes from 'should this be built?' to 'who should build it, how fast, and under whose rules?' The future becomes less a collective choice and more a train we must responsibly board.

That framing appears across opposing camps. Safety advocates may say competition makes development hard to stop, so responsible labs need to stay at the frontier. Accelerationists may say technological development is decentralized and cannot be suppressed, so society should embrace it. Both can end up treating continued development as the fixed variable.

THE MARKETING PARADOX

Utopia and catastrophe can both advertise capability. 'This AI could cure disease' says the technology is powerful. 'This AI could be existentially dangerous' also says the technology is powerful. The same claim can function simultaneously as warning, belief, fundraising narrative and status signal.

THE SYNTHESIS

18. What is actually happening in Silicon Valley?

The cleanest answer is not that one ideology conquered Silicon Valley. It is that several intellectual streams converged in a place where ideas can rapidly become capital, institutions and products.

One stream comes from computing itself: the dream of augmenting human thought. Another comes from Bay Area counterculture: tools, networks, decentralization and self-transformation. Transhumanism adds the proposition that biology is not the final specification. Extropianism adds radical optimism, personal transformation and distrust of centralized limits. Singularity thinking turns machine intelligence into a possible break in history. Rationalism tries to build a culture of better reasoning around these high-stakes questions. Effective altruism adds moral optimization. Longtermism expands the population under consideration to the far future. AI safety asks how to retain meaningful control over systems that may become highly capable. e/acc replies that slowing technological evolution may itself be the greater error. TESCREAL critics ask why a small, wealthy network should get to define the future in these terms at all.

Then comes the special Silicon Valley ingredient: execution. People in this ecosystem are unusually likely to found companies, move billions of dollars, create research institutes, recruit elite technical talent and build actual machines. Philosophy does not remain philosophy for long.

That creates a feedback loop. Futurist ideas motivate ambitious technology. Successful technology makes the old futurism seem less absurd. Increased credibility attracts money and talent. More money and talent produce stronger technology. Stronger technology raises the stakes of the philosophy. The argument then migrates into policy, philanthropy and culture.

The result in 2026 is genuinely strange. Ideas once discussed in tiny mailing lists - human enhancement, machine superintelligence, alignment, posthuman futures, intelligence explosion - now appear, in moderated forms, in official language from companies worth enormous sums and responsible for systems used by hundreds of millions of people. Meanwhile a counterculture of accelerationists mocks the safety culture for being too fearful, and critics of the entire intellectual bundle argue that both sides grant technological elites far too much authority over humanity's future.

None of that tells us who is right about AGI, ASI, consciousness, extinction risk or the future of work. It tells us something more basic: the AI race is not only an engineering competition. It is also a competition between stories about humanity - what we are, what we should become, what risks we should tolerate, and who is entitled to make those choices.

THE SHORTEST POSSIBLE ANSWER

What the heck is happening in Silicon Valley? A seventy-year argument about humans and machines has acquired GPUs, venture capital and corporate legal departments.

IMPORTANT GUARDRAILS

Before you leave: five things this history does NOT prove

- It does not prove that everyone working in AI shares a transhumanist worldview. Most people in AI are doing ordinary research, engineering, product, policy or business work.

- It does not prove that AGI or ASI will happen on any particular timeline - or at all. AGI itself still lacks a universally accepted definition.
- It does not prove that AI-safety concern is merely marketing. The history clearly predates today's commercial boom.
- It does not prove that corporate mission language is free of incentives. Sincere belief, institutional self-interest and marketing usefulness can coexist.
- It does not prove a secret conspiracy. The evidence supports overlapping networks, influence, funding and intellectual inheritance - not a single hidden command structure.

A COMPRESSED CHRONOLOGY

Timeline: from human-computer symbiosis to AGI/ASI

YEAR	WHAT CHANGED
1960	Licklider publishes “Man-Computer Symbiosis,” imagining tightly coupled human-computer thinking.
1962	Engelbart frames computing as a system for augmenting human intellect.
1965	I. J. Good describes the possibility of an ultraintelligent machine improving machine design.
1968	The Whole Earth Catalog popularizes a Bay Area culture of “access to tools,” self-sufficiency and networks.
1972	Peter Singer publishes “Famine, Affluence, and Morality,” a major ancestor of effective-altruist ethics.
1988	Extropy begins publication; the Extropian scene develops around technological self-transformation.
1990	Max More publishes the Extropian Principles.
1993	Vernor Vinge publishes “The Coming Technological Singularity.” Hal Finney writes about electronic cash in Extropy.
1995	Barbrook and Cameron publish “The Californian Ideology,” critiquing the Bay Area fusion of counterculture and market techno-utopianism.
1998	Bostrom and David Pearce found the World Transhumanist Association.
2000	The Singularity Institute for Artificial Intelligence is founded; it later becomes MIRI.
2002	Bostrom publishes his foundational paper on existential risk.
2003	Bostrom publishes “Astronomical Waste,” connecting the far future to moral priority.
2006	Overcoming Bias and the Singularity Summit era help consolidate a rationalist / AI-futurist community.

YEAR	WHAT CHANGED
2009	LessWrong launches as a major rationalist community hub.
2010	At the Singularity Summit, DeepMind founders connect with Peter Thiel through Yudkowsky's network.
2011	The term “effective altruism” gains currency around the Oxford movement; Anthropic does not yet exist.
2014	Bostrom publishes Superintelligence, bringing AI control arguments to a much wider audience.
2015	Musk funds beneficial-AI research through FLI; OpenAI is founded with a humanity-benefit mission.
2021	Anthropic is founded as an AI safety and research company.
2022	Early e/acc writing appears, pushing an explicitly accelerationist response to safety-focused culture.
2023	Andreessen publishes “The Techno-Optimist Manifesto.”
2024	Geburu and Torres publish the TESCREAL paper; Oxford's Future of Humanity Institute closes.
2025	Open Philanthropy becomes Coefficient Giving after directing billions in philanthropy across many causes.
2026	OpenAI discusses a post-AGI transition and personal AGI; Google DeepMind publishes “From AGI to ASI.”

KEEP THIS PAGE

Glossary for normal people who did not spend the 1990s on futurist mailing lists

AI Artificial intelligence: the broad field of building computer systems that perform tasks associated with intelligence.

AGI Artificial general intelligence: a contested term for very general machine capability across many domains, often imagined at or beyond human level.

ASI Artificial superintelligence: hypothetical general machine intelligence substantially beyond human capability.

Alignment The problem of making AI systems reliably behave in accordance with intended goals, constraints and human values.

Transhumanism A family of views supporting the use of technology to overcome or transform important human biological limitations.

Extropianism A 1980s-90s transhumanist philosophy associated especially with Max More, emphasizing self-transformation, technological optimism, intelligence and decentralization.

Singularity A hypothetical future discontinuity in which superhuman machine intelligence or accelerating technological change makes the world beyond it unusually hard to predict.

Rationalism In this context, the LessWrong-associated culture of deliberate reasoning, Bayesian updating, decision theory and reducing cognitive error.

Effective altruism (EA) A movement asking how people can use evidence and reasoning to do as much good as possible with money, careers and other resources.

Existential risk (x-risk) A risk that could destroy humanity or permanently and drastically curtail its long-term potential.

Longtermism A family of moral views giving substantial - and in strong versions, overriding - importance to effects on the long-run future.

e/acc Effective accelerationism: a pro-technological-development movement that emerged partly in opposition to AI-doomer and longtermist caution.

TESCREAL A critical acronym grouping Transhumanism, Extropianism, Singularitarianism, Cosmism, Rationalism, Effective Altruism and Longtermism. Useful as a lens; disputed as a unified category.

Frontier AI The most capable class of AI systems being developed at a given time.

Cryonics Low-temperature preservation after legal death in the hope that future technology might make repair or revival possible; revival of a cryopreserved human has never been demonstrated.

PRIMARY TEXTS, INSTITUTIONAL HISTORIES AND CRITICAL SOURCES

Source notes & further reading

The article prioritizes original texts and institutional histories where possible, then uses reporting or later scholarship to document connections between people and organizations. Links below were checked during research in October 2026. Inclusion does not imply endorsement.

- [1] J. C. R. Licklider, 'Man-Computer Symbiosis' (1960) [Source](#)
- [2] Douglas C. Engelbart, 'Augmenting Human Intellect: A Conceptual Framework' (1962) [Source](#)
- [3] Fred Turner, *From Counterculture to Cyberculture* (University of Chicago Press, 2006) [Source](#)
- [4] Whole Earth Index - archive of Whole Earth publications [Source](#)
- [5] Richard Barbrook & Andy Cameron, 'The Californian Ideology' (1995/1996) [Source](#)
- [6] Nick Bostrom, 'A History of Transhumanist Thought' (2005) [Source](#)
- [7] Max More, 'The Extropian Principles' (Extropy #6, 1990) [Source](#)
- [8] EXTROPY: The Journal of Transhumanist Thought - complete archive, 1988-1996 [Source](#)
- [9] Hal Finney, 'Protecting Privacy with Electronic Cash' (Extropy #10, 1993) [Source](#)
- [10] Alcor Life Extension Foundation, Max More appointed CEO (2010; effective 2011) [Source](#)
- [11] I. J. Good, 'Speculations Concerning the First Ultraintelligent Machine' (1965) [Source](#)
- [12] Vernor Vinge, 'The Coming Technological Singularity' (1993) [Source](#)
- [13] Machine Intelligence Research Institute, institutional history [Source](#)
- [14] LessWrong, 'History of Less Wrong' and 'Rationalist Movement' [Source](#)
- [15] Peter Singer, 'Famine, Affluence, and Morality' (1972) and overview of its influence [Source](#)
- [16] Effective Altruism, 'What is effective altruism?' - movement history [Source](#)
- [17] Nick Bostrom, 'Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards' (2002) [Source](#)
- [18] Nick Bostrom, 'Astronomical Waste' (2003) [Source](#)
- [19] Hilary Greaves & William MacAskill, 'The Case for Strong Longtermism' (2021) [Source](#)

- [20] Coefficient Giving, organizational history (formerly Open Philanthropy) [Source](#)
- [21] WIRED, account of the 2010 Singularity Summit connection between Yudkowsky, Thiel and DeepMind's founders [Source](#)
- [22] Future of Life Institute, Elon Musk's \$10M beneficial-AI research donation (2015) [Source](#)
- [23] OpenAI, 'Introducing OpenAI' (December 2015) [Source](#)
- [24] Anthropic, Series A announcement and founding research aims (2021) [Source](#)
- [25] Anthropic, company mission and Claude's Constitution (current) [Source](#)
- [26] Beff Jezos & bayes, 'Notes on e/acc principles and tenets' (2022) [Source](#)
- [27] WIRED, profile of Guillaume Verdon / 'Based Beff Jezos' and e/acc [Source](#)
- [28] Marc Andreessen, 'The Techno-Optimist Manifesto' (2023) [Source](#)
- [29] Timnit Gebbru & Emile P. Torres, 'The TESCREAL bundle' (First Monday, 2024) [Source](#)
- [30] Future of Humanity Institute, final institutional history and closure notice (2024) [Source](#)
- [31] OpenAI, 'Built to benefit everyone: our plan' (June 2026) [Source](#)
- [32] Google DeepMind, 'From AGI to ASI' (June 2026) [Source](#)
- [33] Google DeepMind, current mission and responsibility/safety pages [Source](#)
- [34] Stanford HAI, 'What is AGI?' [Source](#)
- [35] OpenAI Charter - one influential operational definition of AGI [Source](#)

CLOSING NOTE

The point of learning this history is not to decide in advance that Silicon Valley's futurists are prophets, frauds, heroes or lunatics. It is to notice that the language surrounding AI did not appear from nowhere. Today's arguments carry decades of assumptions about progress, intelligence, markets, human nature and the future. Once those assumptions are visible, we can agree with them, reject them or modify them more intelligently.

And that may be the most democratic response to a technology repeatedly described as civilization-changing: understand the story well enough that the future is not discussed only by the people already inside the room.

FROM EXTROPY TO AGI

Research edition · October 2026