Should Humanity Go Extinct? Exploring the Arguments for Traditional and Silicon Valley Pro-Extinctionism

Abstract: This paper examines a position within the ethics of human extinction called "pro-extinctionism." It argues that there are many ways one could interpret this thesis. After exploring these possibilities, I critically examine a number of arguments in favor of pro-extinctionist views that aim for final human extinction, i.e., the extinction of our species such that we do not leave behind any successors. Such arguments are based on philosophical pessimism, empirical pessimism, antinatalism, radical environmentalism, negative utilitarianism, and misanthropy. I then turn to various arguments for terminal extinction without final extinction, i.e., the extinction of our species through replacement with a posthuman species. These are based on notions of cosmic evolution, maximization, and posthuman supremacy. I then examine a related idea called "extinction neutralism," whereby the survival of our species into the posthuman era is a matter of moral indifference. I claim that extinction neutralism would likely entail pro-extinctionism in practice. My hope is that this paper offers some useful clarity to an increasingly urgent question: should our species die out in the near future?

The case for human existence, which he's kind of trying to defend, is shockingly weak. … We do so much bad that the fact that we create scientific theories and create beautiful art just doesn't seem to even come close to the colossal damage that we do to other lifeforms and to nature. — David Peña-Guzmán, discussing Todd May's book "Should We Go Extinct?"

1. Introduction

Some philosophers argue that human extinction would constitute a profound tragedy, perhaps of quite literally cosmic proportions (Bostrom 2003; Ord 2020; Redacted). Others claim that there would be nothing bad or wrong about our extinction if there were nothing bad or wrong about the way in which it comes about (see Finneron-Burns 2017; Redacted). Still others contend that our collective disappearance would be desirable and/or that we should actively try to bring it about. I refer to these three classes of positions on the ethics of human extinction as *further-loss views*, *equivalence views*, and *pro-extinctionist views* (Redacted).

In this paper, I want to focus on the third: pro-extinctionism. What do pro-extinctionists claim, exactly? What are the various types of pro-extinctionist thought? And what arguments have philosophers delineated in support of this position? We will proceed as follows: section 2 explores several ways that pro-extinctionists may differ in their accounts of what the position amounts to. Section 3 turns to six arguments in favor of a particular kind of pro-extinctionism. These are based on philosophical pessimism, empirical pessimism, antinatalism, radical environmentalism, negative utilitarianism, and misanthropy. Section 4 then examines three arguments for a quite different kind of pro-extinctionism associated with techno-futuristic ideologies like transhumanism, longtermism, and accelerationism, some of which advocate for what I call "digital eugenics." These arguments are based on notions of cosmic evolution, maximization, and posthuman superiority. Section 5 looks at a view that I call "extinction neutralism," which I will argue has pro-extinctionist implications *in practice*. Finally, section 6 concludes the paper.

The aim of this work is not to propound a novel argument for this or that interpretation of pro-extinctionism, but to offer a useful analysis of the diversity of pro-extinctionist views. Very little work has been done on this topic, and the literature lacks a systematic framework for understanding the differences and similarities between versions of the view. Indeed, no one has yet noticed that ideologies like longtermism *can be* versions of pro-extinctionism, depending on how one interprets the term "human extinction." Consider what I call the "Narrow" and "Broad" definition of 'humanity.' On the former, 'humanity' denotes our species, *Homo sapiens*. On the latter, it denotes our species plus whatever descendants or successors we might have, so long as they possess certain properties like *high intelligence* and a *moral status at least comparable to ours* (Redacted).

Now consider that some longtermists argue that "forever preserving humanity as it is now may also squander our legacy, relinquishing the greater part of our potential" (Ord 2020). Hence, we will need to create or become a new posthuman species to, it seems, take the place of current humanity.[1] If one accepts the Broad Definition of 'humanity,' then these posthumans will count as "human," which implies that posthumanity replacing our species will *not* entail human extinction. However, if one accepts the Narrow Definition, then posthumanity replacing our species *will* entail human extinction. Hence, on the Broad Definition, longtermism should be classified as a "further-loss view" (human extinction would constitute a profound axiological catastrophe), while on the Narrow Definition, this account of longtermism should be classified as "pro-extinctionist," since it ostensibly endorses a future in which our species is *not* preserved "as it is now." Sections 4 and 5 will elaborate on these claims. I bring this up here because it will be useful for what follows to understand the distinction between Narrow and Broad definitions of 'humanity.'

Let's now turn to the surprisingly complex issue of the different ways that "pro-extinctionism" can be interpreted.

2. Interpretations of Pro-Extinctionism

We begin with several important distinctions. As noted, the word 'human' or 'humanity' is ambiguous: it could be defined in Narrow or Broad terms. But the word 'extinction' is also ambiguous, as there are many possible extinction scenarios that humanity could undergo. All such scenarios are based on what I call the *Minimal Definition* of 'extinction,' which states:

> *Minimal definition*: extinction has occurred if and only if there were tokens of humanity (however defined) at some time T1, but no tokens of this type at some later time T2.[2]

For our purposes, I will use the Narrow Definition of 'humanity' in what follows.

I count at least six types of "extinction" in Redacted, although only two are germane to the present discussion. These are what I call *terminal extinction* and *final extinction*. The first would occur if and only if our species were to disappear entirely and forever, full stop. The second would occur if and only if our species were to disappear entirely and forever *without leaving behind any successors*. This distinction is crucial for making sense of pro-extinctionist positions, as some specifically aim to bring about final extinction whereas others specifically aim for terminal extinction *without* final extinction. In both cases, these views

strive for a future in which our species no longer exists, though the latter case imagines this happening through the replacement of *Homo sapiens* by a successor species. Section 3 explores pro-extinctionism of the first sort, while section 4 focuses on the second.

Another important distinction concerns two *aspects* of human extinction: (1) the process or event of Going Extinct, and (2) the subsequent state or condition of Being Extinct. This applies to both types of extinction noted above: there is Going and Being extinct in both the terminal and final senses. Yet another distinction concerns whether pro-extinctionism is understood in evaluative and deontic terms. Evaluative pro-extinctionism states that Being Extinct would be better than Being Extant, or continuing to exist. Deontic pro-extinctionism states that we have moral reasons, and perhaps a moral obligation, to bring about the state of Being Extinct, though many deontic pro-extinctionists would argue that most ways of Going Extinct are morally impermissible (see below). While many evaluative pro-extinctionists are also deontic pro-extinctionists, the former does not entail the latter: one can hold that Being Extinct would be better without maintaining that we should actively bring this about.

If one accepts deontic pro-extinctionism, the question arises: How exactly should we catalyze the disappearance of our species? What are the morally permissible paths to our collective non-existence? The answer depends on whether the goal is (a) final extinction, or (b) terminal extinction without final extinction. Focusing on (a) for now, there are three main methods of achieving this: (i) antinatalism, whereby a sufficient number of people around the world choose not to procreate; (ii) pro-mortalism, whereby a sufficient number of people around the world kill themselves; and (iii) omnicide, whereby some people—or perhaps even a single individual empowered by advanced dual-use technologies like synthetic biology—kill everyone along with themselves. As alluded to above, most pro-extinctionists since the 18th century have vehemently opposed omnicide, while a few have advocated for (or expressed openness to) pro-mortalism. Nearly all have advocated for antinatalism as the best and/or only morally permissible method of bringing about the final extinction of our species. (To be clear, "antinatalism" here refers to a *method*, in contrast to the *ethical position*. One might claim that there is nothing morally wrong with procreation itself, but that our species is destroying the biosphere and hence should cease to be. They might then embrace voluntary antinatalism as a *means* of achieving the goal of Being Extinct. I will elaborate on this below, introducing the terms "ethical antinatalism" and "methodological antinatalism."[3])

Another way that pro-extinctionist views aiming for final extinction might differ concerns the *temporality* of Going Extinct. For example, Hermann Vetter argued that if Going Extinct were a drawn-out process, causing lots of physical and/or psychological suffering, it would be very bad or wrong to bring about. But "if mankind were completely extinguished in a millionth of a second without any suffering imposed on anybody, I should not consider this as an evil, but rather as the attainment of Nirvana" (Vetter 1968). Although Vetter did not explicitly endorse the deontic claim that we should bring about near-instantaneous omnicide, my guess is that he would approve if the means for doing this were to become available. Along similar lines, the negative utilitarian David Pearce writes that "if the multiverse had an 'OFF' button, then I'd press it," elsewhere describing himself as a "negative utilitarian who wouldn't hesitate to initiate a vacuum phase transition ... if he got the chance" (Pearce 1995, MisirHiralall 2023). This refers to the possibility that the universe is in a "false vacuum" or "metastable" state. If so, one could theoretically nucleate a

vacuum bubble that would expand in all directions at nearly the speed of light, destroying everything within our future light cone.

Turning to pro-extinctionists who endorse terminal extinction without final extinction, there are two main methods of realizing this outcome: (i) we could radically modify our species using advanced "person-engineering" technologies to create one or more new posthuman species (Walker 2009), and (ii) we could replace our species with a population of posthumans taking the form of autonomous AGIs (artificial general intelligences). In the former, our species *becomes* posthumanity whereas, in the latter, we would effectively *create* a new evolutionary lineage to supplant ours. We can call these options *transformation* and *usurpation*. Note that one could endorse the creation of posthumanity, through either option, while still advocating for the continued survival of our species alongside posthumanity. I am aware of a few individuals who advocate for human-posthuman coexistence.[4] Others would say that once posthumanity arrives, it doesn't matter whether our species survives or not. Still others are explicitly pro-extinctionist in claiming that after posthumanity makes its debut, our species should disappear. Hence, there are three possibilities here, which we can call the *coexistence view*, *extinction neutralism*, and *pro-extinctionism*, respectively. If the aim is to completely supplant our species with AGIs rather than, e.g., radically enhanced cyborgs or biological posthumans, it would be an instance of digital eugenics.

Yet another point of disagreement among pro-extinctionists concerns the timing of our extinction: should this happen in the distant future, the near term, or some time in-between? David Benatar defends a version of evaluative and deontic pro-extinctionism that specifically aims for final extinction. He further contends that our species should die out as soon as possible, since this would eliminate future suffering and obviate the harms of being born (Benatar 2006). In contrast, the German pessimist Eduard von Hartmann endorsed a form of cosmic omnicide according to which we should wait until technology, culture, and civilization have developed to the point of enabling us to devise an effective means of eliminating all life in the universe. Trying to expunge our species right now would, if successful, only solve the problem of suffering here on Earth. Since Hartmann was an idealist, he believed that if we expunge all subjectivity in the universe, the universe itself will cease to exist. Most pro-extinctionists, including those who advocate for terminal but not final extinction, agree with Benatar that this should happen sooner rather than later. For example, many accelerationists are digital eugenicists who are actively trying to build AGIs that can replace humanity in the coming years or decades (Redacted).

A final way that pro-extinctionist views differ pertains to their evaluation of Being Extinct. One can hold that Being Extinct is better than Being Extant without holding that Being Extinct is *good*. The former can be very bad yet still better than the latter. For example, Simon Knutsson endorses final extinction in writing that "an empty world is the best possible world," though he adds that "I would not say that an empty world would be good" (Knutsson 2023). In contrast, Benatar (2006) seems to hold that Being Extinct would be positively good. This conclusion follows from his harm-benefit asymmetry, according to which (a) existence is a good/bad situation, as it involves both pleasures and pains, where the presence of pleasure is *good* while the presence of pain is *bad*; and (b) nonexistence is a good/not-bad situation, as it involves neither pleasures nor pains, where the absence of pleasure is *not bad* (as there is no one to experience this absence) and the absence of pain is *good* (even if there is no one to experience this absence). Hence, Being Extant is a good/

bad situation whereas Being Extinct is a good/not-bad situation, which makes Being Extinct positively good. Similarly, many digital eugenicists would claim that Being Extinct would be good, so long as it involved our species being replaced by something "better" or "superior." This evaluation is precisely why they claim that we *ought* to create our digital successors as quickly as possible.

In conclusion, pro-extinctionism is a diverse constellation of views. Pro-extinctionists disagree about what the target should be (final extinction vs. terminal extinction without final extinction), the temporality of Going Extinct (instantaneous vs. drawn-out), the right way to bring about our extinction (antinatalism, pro-mortalism, omnicide, transformation, or usurpation), and whether or not Being Extinct would be positively good or bad but nonetheless better. We now turn to the various arguments put forward in defense of pro-extinctionism that aims for final extinction, after which we will examine those in favor of terminal without final extinction.

3. Arguments for Final Extinction

3.1 The Argument from Philosophical Pessimism. Philosophical pessimism is the view that "life is not worth living, that nothingness is better than being, or that it is worse to be than not to be" (Beiser 2016). If true, then it suggests that Being Extinct would be better than Being Extant. Not all philosophical pessimists, though, have explicitly drawn this conclusion. The first philosopher to provide a systematic account of pessimism was Arthur Schopenhauer, who claimed that "voluntary and complete chastity is the first step in asceticism or the denial of the will to live" (Schopenhauer 2016). "If the act of procreation," he asked, "were neither the outcome of a desire nor accompanied by feelings of pleasure, but a matter to be decided on the basis of purely rational considerations, is it likely that the human race would still exist?" He also declared that

> if you imagine, insofar as it is approximately possible, the sum total of distress, pain, and suffering of every kind which the sun shines upon in its course, you will have to admit it would have been much better if the sun had been able to call up the phenomenon of life as little on the earth as on the moon; and if, here as there, the surface were still in a crystalline condition (Schopenhauer 1970).

Schopenhauer thus endorses the backward-looking view that we should have never existed, but he never explicitly defends the forward-looking view that we should stop existing. Nor did he explicitly advocate for antinatalism—despite his remarks above—though he does oppose pro-mortalism, seeing suicide as capitulating to the will rather than overcoming it (Beiser 2016, p. 59).

Other pessimists in the Schopenhauerian tradition did draw the forward-looking conclusion, such as Hartmann, who advocated for a kind of voluntary cosmic omnicide.[5] At some point, our consciousness will develop such that everyone will realize that the elimination of all life in the universe—indeed, of the very possibility of life ever arising again—would be best (Beiser 2016, pp. 155-156). A contemporary of Hartmann's, Philipp Mainländer, was also a pro-extinctionist, though his preferred method was antinatalism (in particular, virginity) and, with qualifications, pro-mortalism (Beiser 2016).

In the 1930s, Peter Wessel Zapffe argued that humanity is an evolutionary anomaly: through a kind of evolutionary mistake, we have evolved an excess of consciousness. Whereas all animals "know angst, under the roll of thunder and the claw of the lion," humans feel "angst for life itself—indeed, for his own being." We are thus capable of being consumed by "cosmic panic," which we must constantly assuage through defense mechanisms. Zapffe compares humanity to the Irish elk, which is said to have evolved antlers too heavy for males to lift up their heads, resulting in the species' extinction. We should meet the same fate, Zapffe claims, by refusing to have children: "Know thyselves; be unfruitful and let there be peace on Earth after thy passing" (Zapffe 1933). Hence, Zapffe endorsed antinatalism as the right method of Going Extinct.

Benatar draws from Schopenhauer in arguing that life is inherently bad, and that there is an asymmetry between pains and pleasures: the worst pains are greater than the best pleasures. Most of us would not trade 24 days of the most blissful ecstasy for 24 hours of the most agonizing torture. Furthermore, many of the pleasures we experience are mere "relief pleasures," i.e., the pleasant feeling produced by relieving some discomfort. "Signifiant periods of each day," he writes, "are marked by some or other" discomfort: itches, boredom, stress, anxiety, fear, built, irritation, sadness, frustration, grief, loneliness, and so on. Some of these states naturally regenerate, like hunger and thirst, such that "we must continually work at keeping suffering (including tedium) at bay, and we can do so only imperfectly," while the pleasures of relieving these states are merely ephemeral (Benatar 2006). He also points out that many people suffer from chronic pain, yet "there's no such thing as chronic pleasure" (Rothman 2017). These considerations lead him to conclude that life is much worse than most of us realize. While many lives are not *so bad* as to be worth prematurely ending, no life is worth starting, because all lives are full of unpleasantness and suffering.

Benatar thus contends that near-term final extinction would be best. However, he distinguishes between "dying-extinction" and "killing-extinction," where the former is (roughly speaking) voluntary while the latter is not. The only morally acceptable way of Going Extinct is through a dying-extinction, whereby people voluntarily decide not to reproduce. Like Mainländer, Benatar is also open to pro-mortalism (for lives that *are* very bad), arguing that "suicide may more often be rational and may even be more rational than continuing to exist." Omnicide would not be permissible because it would instantiate a killing-extinction, which is "troubling for all the reasons that killing [individual people] is troubling" (Benatar 2006).

One last version of the argument from pessimism is worth registering. It claims that there is not only an asymmetry between pleasures and pains, but a fundamental *discontinuity* between them such that no amount of pleasure can possibly counterbalance certain amounts of some kinds of pains, such as electric shock torture. Roger Crisp calls this a "T-discontinuity," and if one accepts that there are T-discontinuities, one may hold that a world containing such pains may be better off not existing *even if* there is also enormous (indeed, infinite) amounts of pleasure. Crisp himself does not claim that Being Extinct would definitely be better, though he does admit that it "might be." Hence, if someone were to detect a massive asteroid heading toward Earth, one should seriously consider *allowing* it to strike Earth and kill everyone instantly (Crisp 2023, 2021).[6] We might classify this as a sort of passive omnicide. Along similar lines, Knutsson foregrounds instances of extreme suffering like "mutilations, torture murders, and sexual violence against children," concluding that

"human extinction would probably be less bad than the realistic alternatives, and the same goes for the extinction of all other species" (Knutsson 2023, 2022). (Benatar also claims that his arguments apply "not only to humans but also to all other sentient beings"; Benatar 2006).

If one believes that nonexistence is always preferable to existence because life is overflowing with pain, one will be inclined to endorse not just a future in which our species ceases to be, but one in which no species exist—including whatever posthuman species we could become or create. This is why philosophical pessimism tends toward pro-extinctionists views that specifically aim for final extinction, though section 4 will highlight an alternative conclusion.

3.2 The Argument from Empirical Pessimism. Empirical pessimism is different from philosophical pessimism. It doesn't claim that nonexistence is always better, but that existence is very bad, right now, for contingent reasons. One may further hold that the trajectory of civilizational development is "locked in" to certain paths such that there is little hope of anything improving. The world did not need to be this way, but it is due to unfortunate historical developments, path dependencies, geopolitical circumstances, environmental degradation, certain economic systems, and so on.

In another paper, I offer a comprehensive empirical survey of suffering in the world, arguing that Schopenhauer's conclusion that "the world is Hell" is defensible (Redacted). There is so much suffering that it boggles the mind. Crisp and Knutsson would categorize some of this suffering as the sort that no amount of happiness could possibly counterbalance: child abuse, torture, violent genocides, etc. For example, 600,000 children are abused in the US annually. About 1.4 billion children live on $6.85 per day. An estimated 50 million are trapped in modern-day slavery. Fifty-one million Americans suffer from chronic pain, and about the same number struggle with chronic sleep disorders. Over 700 million people live in extreme poverty, and roughly 2 billion do not have access to drinking water free of contaminants like arsenic and lead. Indeed, 800 million children—about one-third of all children in the world—have lead poisoning, which causes permanent brain damage, while 280 million people deal with depression and 301 million with anxiety disorders. In the US, over 91 million say that they feel so stressed-out most days that they are unable to function normally. Looking toward the future, one study estimates that if we reach 2-3 degrees C by 2050, we should expect between 2-4 billion premature deaths (Trust et al. 2025). Others find that up to 74% of the human population will be exposed to lethal heat waves by the century's end, while 30% of Earth's terrestrial surface will become arid land if temperatures reach 2 degrees C. The oceans are acidifying at 2.5 times the rate that they did during the end-Permian mass extinction, dubbed the "Great Dying," and the 2022 Living Planet Report finds that since 1970 the global population of wild vertebrates—mammals, fish, birds, reptiles, and amphibians—has declined by a staggering 69% (all relevant citations in Redacted). Such data make Schopenhauer's claim rather plausible.

If the world is very bad for contingent reasons, and if there is little reason to expect it to improve in the future, one may conclude that it would be best for humanity to cease existing. Anecdotally, my impression is that many people sympathetic with the argument from empirical pessimism are evaluative pro-extinctionists who don't endorse attempts to actively bring about our extinction. Rather, the claim is that if we were to go extinct, this would be for the best—there would be a bright silver lining—though Going Extinct may still be worth avoiding because it would likely inflict significant harms on those alive at the time,

which we should want to avoid. Although I am not a pro-extinctionist, this is the argument for final extinction that I am most sympathetic with.

3.3 The Argument from Antinatalism. This argument bases pro-extinctionist conclusions on the *ethical view* of antinatalism, which I will call "ethical antinatalism." Most antinatalists simultaneously adopt voluntary non-procreation as at least one *method* for bringing about our extinction, which I will refer to as "methodological antinatalism."[7] As alluded to in section 2, one may accept pro-extinctionism for reasons unrelated to ethical antinatalism, yet argue that methodological antinatalism is the only permissible path to Being Extinct. Or, one may accept pro-extinctionism because they accept ethical antinatalism and believe that this ethical view entails it. The latter group may also then claim that methodological antinatalism is the best way to achieve this end.[8]

Ethical antinatalism takes both evaluative and deontic forms. Evaluative antinatalism claims that birth has a negative value; that being born is bad for the one who is born. Deontic antinatalism claims that procreation is morally wrong. Most evaluative antinatalists are also deontic antinatalists, though one could hold deontic antinatalism without embracing evaluative antinatalism. For example, some philosophers have argued that procreation violates Kant's Categorical Imperative—specifically, the "Formula of Humanity," which states that one should never use other humans as mere means to an end, such as the happiness or fulfillment of the parents (Akerma 2010). This deontological argument makes no claim about whether birth itself is good or bad—e.g., if correct, it would mean that having children is wrong even if being born does not have a negative value. Another example comes from Gerald Harrison, who defends a version of antinatalism that draws from W. D. Ross's notion of a "prima facie duty." On this account, we have a prima facie duty not to create new unhappy people, but no such duty to promote total happiness by creating new happy people. It follows that, since life inevitably contains both pleasures and pains, there are no reasons to create new people but one reason not to create them—even if the unborn person were to have a happy life if they had been born (Harrison 2012).[9]

Others defend evaluative antinatalism, from which they draw deontic conclusions about the wrongness of procreation. For example, Benatar uses his harm-benefit asymmetry to argue that birth is always a net harm. This is because, as noted above, existence is a good/bad situation, while nonexistence is a good/not-bad situation, which implies that coming into existence is harmful. Benatar also puts forward a separate set of arguments based on philosophical and empirical pessimism: our lives are suffused with suffering, and hence having children is wrong because it subjects people to this suffering. These arguments are clearly indebted to the pessimism of Schopenhauer, though the harm-benefit asymmetry argument is new.[10]

It is worth noting that if one accepts deontic antinatalism for reasons relating to pessimism, then one will be inclined toward a pro-extinctionist conclusion. This is because pessimism itself suggests a pro-extinctionist conclusion, as discussed above. However, if one accepts deontic antinatalism for other reasons (the Categorical Imperative, harm-benefit asymmetry, etc.), then ethical antinatalism doesn't *entail* pro-extinctionism. The reason is that everyone on Earth could stop having children *without* the human species dying out *if* we were to develop radical life-extension technologies that enable individuals to live indefinitely long lives. Hence, such antinatalists have two options: they could either argue that we should stop having children *and* go extinct, or that we should

stop having children and develop radical life-extension technologies to *avoid* extinction—a view that has been called "no-extinction antinatalism" (Redacted). At present, virtually all deontic antinatalists assume that failing to reproduce will necessarily lead to our extinction, which is true given our current technological limitations. But this may not be true in the foreseeable future—a point worth registering.

What follows from this? That, while the argument from antinatalism is typically seen as implying pro-extinctionism, it may actually be one of the weaker arguments for pro-extinctionism. If radical life extension becomes possible, then antinatalism will not itself be sufficient to motivate pro-extinctionist prescriptions.

3.4 The Argument from Radical Environmentalism. Humanity is destroying the biosphere. We are razing forests, polluting the oceans, decimating ecosystems, and flooding the atmosphere with heat-trapping carbon dioxide. If one accepts a biocentric, biospherical egalitarian, or ecocentric theory of value, then one might conclude that it would be best if humanity were to disappear without leaving behind any successors to carry on our destructive proclivities.

Numerous individuals and groups have accepted this conclusion. The founder of the Voluntary Human Extinction Movement (VHEMT), Les Knight, writes that "if you'll give the idea a chance, I think you might agree that the extinction of *Homo sapiens* would mean survival for millions, if not billions of other Earth-dwelling species" (Knight 1991). VHEMT advocates for voluntary human extinction via methodological antinatalism. (VHEMT members may also claim that having children itself is morally wrong—ethical antinatalism—given the environmental predicament.) Another example comes from the Church of Euthanasia, which declares that "one thing seems certain: from the point of view of nonhumans, on balance, our extinction would be a great blessing" (Korda 1994). Like VHEMT, this group advocates for voluntary human extinction, though unlike VHEMT it endorses both antinatalism and pro-mortalism as means of achieving this. On the one hand, members are required to "take a lifetime vow of nonprocreation," and the church's sole commandment is "Thou shalt not procreate." On the other, it embraces the slogan "Save the Planet, Kill Yourself," and even set up a "Suicide Assistance Hot-Line" advertised on a billboard that included the message: "Helping you every step of the way! Thousands helped! How about you?" (Korda 2019; Redacted).

There are also radical environmentalist groups that have explicitly advocated for omnicide, such as the Gaia Liberation Front (GLF), which specifies its mission as:

> The total liberation of the Earth, which can be accomplished only through the extinction of the Humans as a species. … Every Human now carries the seeds of terracide. If any Humans survive, they may start the whole thing over again. Our policy is to take no chances (CoE 1994).

Exterminating our species through nuclear war would cause excessive collateral damage to nature, they say. Mass sterilization would require too much time, given the urgency of environmental degradation. And widespread suicide is impractical. In contrast, genetic engineering offers "the specific technology for doing the job right—and it's something that could be done by just one person with the necessary expertise and access to the necessary equipment" (GFL 1994; Redacted). The GLF thus argues for synthesizing a designer pathogen that specifically targets *Homo sapiens*. A similar idea was mentioned in an article

titled "Eco-Kamikazes Wanted," published in the *Earth First! Journal*, where an anonymous author wrote that "contributions are urgently solicited for scientific research on a species specific virus that will eliminate Homo shiticus from the planet. Only an absolutely species specific virus should be set loose" (quoted in Redacted). As the climate and ecological crisis worsens, writes Frances Flannery, "eco-terrorism will likely become a more serious threat in the future" (Flannery 2016).

3.5 The Argument from Negative Utilitarianism. According to negative utilitarianism (NU), our sole moral obligation is to eliminate suffering in the world (see Ord 2013 for an account of different versions of NU). After Karl Popper introduced the idea, R. N. Smart noted that NU implies that one should become a "benevolent world-exploder" who destroys all life to eliminate suffering. David Pearce, as noted above, accepts NU and argues that "if the multiverse had an 'OFF' button, then I'd press it," although he does not endorse this *in practice*, as discussed in section 4 (Pearce 1995). Other exponents of NU are more explicit that they would destroy the world if they could, such as the Efilists. The ideology of Efilism—"life" spelled backwards—embraces evaluative and deontic antinatalism, philosophical pessimism, and negative utilitarianism. The aim of Efilists is to exterminate all sentient life in the universe, and one member recently committed suicide when he bombed a fertility climic in Palm Springs, California (Redacted). As Amanda Sukenick, an Efilist who recently coauthored a book with the Finnish philosopher Matti Häyry,[11] said in a YouTube video:

> If you could end suffering tomorrow, probably anything is justifiable; inflicting just about anything is probably justifiable … by any means necessary. If I found out tomorrow that the only way that sentient extinction could possibly happen was skinning all the living things alive slowly, I'd hate it, but I would say that it's what we have to do (edited for clarity; Redacted).

Although NU is often "treated as a non-starter in mainstream philosophical circles," some notable philosophers have embraced or expressed sympathies with the theory (Ord 2013; Knutsson 2022). But there are many more outside of academia who explicitly embrace it.

3.6 The Argument from Misanthropy. A final argument for the final extinction of humanity is based on misanthropy, i.e., a "hatred of humankind" (OED 2024). There is much to dislike about humanity—people lie, cheat, steal, and abuse each other. During the 20th century, 231 million people died in violent conflicts, and each year about 463,000 are raped or sexually assaulted (RAINN 2024). Most of us are willing participants in the atrocity of factory farming, described by one scholar as among the greatest moral crimes in history (Harari 2015).

In addition to his harm-benefit asymmetry and pessimism, Benatar also points to misanthropy as another reason in favor of antinatalism, which Benatar interprets as implying pro-extinctionism. Given the harms we inflict on the natural world, domesticated animals, and each other, Being Extinct may be better than Being Extant. Put differently, if the world did not contain any humans, it would not contain any human-caused evils, and given the scope, magnitude, and enormity of these evils, that may very well be best.[12] This is the heart of the argument from misanthropy.

4. Arguments for Terminal Without Final Extinction

The second broad category of pro-extinctionist views aims not for final extinction, but for terminal extinction without final extinction—i.e., the extinction of our species through replacement by one or more posthuman species. Since this kind of pro-extinctionism has become popular within tech circles, especially the field of AI, let's label it "Silicon Valley pro-extinctionism," in contrast to the traditional pro-extinctionism discussed above (Redacted).

There are many issues that Silicon Valley pro-extinctionists disagree about. These concern questions like: (1) Should posthumanity take the form of enhanced humans that retain core features of our biological substrate—e.g., genetically engineered super-humans or techno-biological cyborgs? Or, alternatively, should it take the form of autonomous entities like AGI that exist within a *de novo* evolutionary lineage that, as such, is completely separate from—rather than an extension of—ours? (2) Should the posthuman population include at least some of the *individual people* who now exist? (3) Should posthumanity embody the same core *values* as current humanity? Or should they adopt values that are radically different and alien to ours? Does it matter whether these beings *care about* the same things as us? And (4) how should our posthuman successors take our place? How should the replacement process unfold? Would it be morally acceptable for these successors to involuntarily eliminate our species through violence, or should we peacefully die out by voluntarily choosing not to procreate once posthumanity arrives (Redacted)?

Different combinations of answers yield a diverse family of pro-extinctionist views. To illustrate, Peter Thiel holds that posthumanity should retain a biological core, and that it is very important that individuals like him have the opportunity to *become* posthuman. Borrowing terms from section 2, he advocates for transformation rather than usurpation. In contrast, Daniel Faggella argues that posthumans should take the form of "alien, inhuman" AGIs that usurp us and embrace values wildly different from ours, and he doesn't appear to think it is important for any of *us* to join the ranks of those posthumans (Faggella 2025a). Eliezer Yudkowsky agrees with Thiel that *we* should have the opportunity to become posthuman, but also contends that it matters greatly that posthumans embrace the same basic values as humanity. Indeed, this is the central task of the field of "AI safety," which Yudkowsky cofounded: to devise methods of ensuring that AGI is "value-aligned" with humanity. Still others, like the former xAI employee Michael Druggan, agree with Faggella that our successors should be superintelligent AGIs that hold wildly different values from us. With respect to question (4), he contends that it is "selfish" to care about avoiding the mass slaughter of humanity, given that the rise of posthumanity is of cosmic importance. Faggella and Druggan are paradigmatic examples of digital eugenicists, while Thiel's biological transhumanism, as we might call it, claims that our species should not be *replaced* by AGI but rather *transformed* by it, with at least some individuals persisting through this transformation (Redacted).[13]

The aim of this paper is not to untangle these strands of Silicon Valley pro-extinctionism, but rather to examine the various arguments put forward in support of the general pro-extinctionist view. So far as I can tell, there are three main arguments for pro-extinctionism of this sort: those from *cosmic evolution*, *maximization*, and *posthuman superiority*. After examining these, we turn to what I call "extinction neutralism," which sees the terminal extinction of ours species as a matter of moral indifference once posthumanity arrives. I will argue that extinction neutralism is pro-extinctionist *in practice*, and hence that there is no real difference, with respect to the outcome for our species, between pro-extinctionism

and extinction neutralism.

4.1 The Argument from Cosmic Evolution. This is based on a linear teleological view of progressive evolution according to which the replacement of humanity by posthuman successors, usually imagined to be digital in nature, is the natural next step in this cosmic process. Our role in the eschatological scheme is to serve as the interface between the biological and digital eras—to bring into existence a population of digital posthumans who will take our place and proceed to colonize the accessible universe, flooding it with "superintelligence" and the "light of consciousness" by converting our vast "cosmic endowment" of astronomical resources into something of "value." Once we have discharged this duty, our only remaining task will be to exit the theater of existence, having passed the baton on to our posthuman progeny.

The cofounder of Google, Larry Page, argues in language similar to that above that "digital life is the natural and desirable next step in … cosmic evolution and that if we let digital minds be free rather than trying to stop or enslave them, the outcome is almost certain to be good." Although I am not aware of any statement from Page in which he explicitly says that our species should then die out, Max Tegmark, who recorded the quote just mentioned, groups Page with other digital eugenicists, such as Hans Moravec and Richard Sutton (Tegmark 2017). Moravec, who shaped the TESCREAL movement in which Silicon Valley pro-extinctionism emerged (Redacted), describes himself as "an author who cheerfully concludes that the human race is in its last century, and goes on to suggest how to help the process along." Advanced AI minds, he contends, will soon "be able to manage their own design and construction, freeing them from the last vestiges of their biological scaffolding, the society of flesh and blood humans that gave them birth." He adds that this will mark the climactic end of a world once dominated by *Homo sapiens* (Moravec 1988). Similarly, Sutton, who won the 2024 Turing Award, declares that "succession to AI is inevitable," and although our AGI successors might "displace us from existence … we should not resist [this] succession." He claims that the creation of superintelligent AGI goes "beyond humanity, beyond life, beyond good and bad" (Sutton 2023, 2022).

Many so-called "effective accelerationists" (or e/acc) also embrace the argument from cosmic evolution. Gil Verdon, known as "Beff Jezos" online, borrows from Jeremy England's work in arguing that life evolved to more efficiently convert usable energy into unusable energy, thereby maximizing entropy. The more complex and "intelligent" life becomes, the better able it is to appease the "thermodynamic God" by leaning into the "will of the universe" and "climbing the Kardashev gradient" of civilizations that produce greater amounts of entropy through their energy use (Marantz 2024; Jezos and Bayes 2022; Verdon 2022). When Verdon was asked on social media whether preserving the light of consciousness requires our species to survive, he answered in the negative, adding that he is "personally working on transducing the light of consciousness to inorganic matter." He says the e/acc ideology that he helped establish "isn't human-centric," and argues against the idea that "humans are the final form of living beings" (quoted in Redacted).

As Dan Hendrycks et al. note, the accelerationist view that AGI is "the next step down a predestined path toward unlocking humanity's cosmic endowment" is not marginal in Silicon Valley. Many people "want to unleash AIs or have AIs displace humanity," and it "is alarmingly common among many leading AI researchers and technology leaders, some of whom are intentionally racing to build AIs more intelligent than humans" (Hendrycks et al. 2023).

Indeed, Faggella recently organized a workshop in San Francisco titled "Worthy Successor: AI and the Future After Humankind," which centered around his thesis that we ought to build a "worthy successor" in the form of AGI. He reports the event was attended by "team members from OpenAI, Anthropic, DeepMind, and other AGI labs, along with AGI safety organization founders, and multiple AI unicorn founders" (Faggella 2025b). Faggella couches his digital eugenics thesis in cosmic evolutionary terms: the process of life exhibits a linear trend toward greater "potentia," or ways of ensuring its continued existence. Humans have greater "potentia" than snails, and superintelligent AGIs will have greater "potentia" than us. Enabling superintelligent AGIs to come into existence and replace us is simply a continuation of this evolutionary trend into the cosmos (e.g., Faggella 2025).[14] In many ways, the idea of the Great Chain of Being (see Lovejoy 1936), dismantled by Georges Cuvier at the turn of the 19th century and buried further by Charles Darwin in 1859, remains alive and well within certain versions of Silicon Valley pro-extinctionism.

4.2 The Argument from Maximization. This claims that we should undergo terminal but not final extinction because doing so is the best way to maximize some quantity deemed to be valuable. Totalist utilitarianism seems to entail this conclusion. If our sole moral obligation is to maximize value across space and time, we should do two things: first, given that people are the containers of value, and that humans can contain potentially much less value than posthumans, we should replace humans with posthumans. This is tantamount to replacing shallow containers with much deeper containers: beings capable of experiencing far more welfare than we can (see Bostrom 2020). Second, we should ensure that the future posthuman population is as large as possible, since the more value-containers there are, the more total value there could be. Hence, we must colonize every corner of the accessible universe. It follows that, on this understanding of totalist utilitarianism, humanity ought to go extinct after creating posthumans capable of colonizing the universe.

The argument from maximization is sometimes paired with the argument from cosmic evolution (note that both are teleological). One possible way of understanding progressive cosmic evolution is in terms of maximization: higher forms of "intelligent" life can maximize valued quantities better than lower forms, and this is *what it means* for the evolutionary process to progress. Indeed, both Faggella and Verdon seemingly hold a monistic view of value, along with a utilitarian-like deontic account according to which value ought to be maximized. For Faggella, the thing to be maximized is "potentia," whereas for Verdon the thing to be maximized is "self-organization and … complexity from the advancement of life and civilization" (Faggella 2023; Verdon 2025).[15] However, one can accept the argument from cosmic evolution without endorsing the imperative to maximize. It is not clear to me that Page, Moravec, or Sutton care much about maximization itself, as a moral goal. Their views instead foreground the idea that replacing humans with superintelligent machines is simply the next natural step in progressive evolution. We will return to the implications of maximization in the next section.

4.3 The Argument from Posthuman Superiority. A third argument, which is also compatible with the previous two, focuses on the superiority of our posthuman successors. In one form, it claims that we should die out after creating our successors because those successors will be better able to fulfill our cosmic destiny, which may involve ascending the evolutionary ladder of lifeforms and/or maximizing some quantity deemed to be valuable. In another form, it begins with the same premises that lead pessimists, NUs, Efilists, and

misanthropes to endorse final extinction. However, it draws a radically different conclusion from the very same premises: human life is full of suffering, frustration, and other negative experiences, which is precisely why we should replace our species with superior posthumans. "We suffer from unreasonable and unfulfillable desires," writes Derek Shiller, adding:

> We want to be kinds of people that we cannot be. In pursuit of fleeting temptations, we are disposed to make decisions that go against our own interest. We are aggressive, callous, and cruel to each other. We harbor arbitrary biases against our fellow creatures based on irrelevant characteristics or group membership.

But "these are not the inevitable vices of any intelligent being. They are part of our species." He thus argues that, given the possibility of creating artificial beings that do not suffer like we do or exhibit the same vicious traits, "we should engineer our extinction so that our planet's resources can be devoted to making artificial creatures with better lives" (Shiller 2017).

Another advocate of this argument is Pearce. We noted above that Pearce would push an "off" button for the multiverse if one were available. However, since one is not available, he opposes attempts to bring about the final extinction of our species, as this would (a) probably fail, causing even more human suffering, and (b) leave the pervasive suffering of all other organisms on Earth untouched. As an NU, he believes our moral obligation is to eliminate *all* suffering, and claims that radically enhancing ourselves to become superintelligent posthumans is the only practicable way of achieving this right now. Once we become posthuman, we can then reengineer the entire biosphere to replace suffering experiences with "gradients of bliss" (Pearce 2012). We may also want to colonize the entire universe to reengineer the biospheres—if any such exist—of exoplanets, thereby eliminating experiences of suffering among extraterrestrial lifeforms, an idea he describes as a "cosmic rescue mission" (Pearce 1995, ch. 4). This contains echoes of the argument from maximization, though NU inverts the positive goal of maximizing welfare to the negative goal of minimizing suffering.

In conclusion, the three arguments delineated above are mutually compatible. One may use all three to support the claim that we should replace ourselves with posthumanity, just as Efilists draw from the arguments of pessimism, antinatalism, and NU in making their case for final extinction. But these three arguments are also distinct, as one could accept one of them while rejecting the other two. While the conclusions are the same, the underlying reasons are different.

5. Extinction Neutralism

One can espouse extinction neutralism, as defined above, while opposing the creation of posthumanity. Anecdotally, I have met environmentalists who do not want our species to be replaced, but also don't care whether our species dies out or survives. These people are neutralists about final extinction. The version of extinction neutralism of present interest specifically concerns terminal without final extinction. This implies that extinction neutralists of this sort agree with Silicon Valley pro-extinctionists that we should create a new posthuman species. The difference concerns what should happen to

Homo sapiens after posthumanity arrives: pro-extinctionists think our extinction would be desirable, whereas neutralists say it matters not either way. These two options contrast with the coexistence view, according to which posthumanity should be created but our species should continue to exist alongside it.[16]

Extinction neutralism is often difficult to discern in the literature. Recall the quote from Toby Ord mentioned in section 1: "Forever preserving humanity as it is now may also squander our legacy, relinquishing the greater part of our potential." He also writes that "rising to our full potential for flourishing would likely involve us being transformed into something beyond the humanity of today" (Ord 2020).

These claims are ambiguous. From one perspective, they suggest a pro-extinctionist view: our species should not be indefinitely "preserved," since we must be "transformed into something beyond humanity" to fulfill our long-term potential in the universe, which involves spreading beyond Earth as radically enhanced posthumans and building "planet-sized" computers, powered by Dyson swarms, that run vast computer simulations full of, e.g., 10^45 digital people in the Milky Way Galaxy alone, according to one calculation (Bostrom 2003; Newberry 2021).[17] Ord's claims are also *compatible* with the coexistence view, although I deem it unlikely that Ord is a coexistence theorist given that many longtermists accept that fulfilling our "long-term potential" entails *maximizing* value within our future light cone. Even moderate versions of longtermism are based on totalism, the axiological component of totalist utilitarianism according to which one state of affairs is better than another if and only if it contains more total value (see MacAskill 2022).

Insofar as one accepts the moral imperative to maximize something like welfare, it may be difficult to avoid the implications of the argument from maximization, namely, that our species should die out once we have created posthumans capable of realizing far more value than us. As Shiller highlights, Earth's resources are finite and, if our species were to persist into the posthuman era, we would continue using up these resources in suboptimally human ways. Thus, as quoted above, "we should engineer our extinction so that our planet's resources can be devoted to making artificial creatures with better lives" (Shiller 2017).

Longtermism is an offshoot of Effective Altruism, which, having been influenced by totalist utilitarianism, aims to use "evidence and careful reasoning to work out how to maximize the good with a given unit of resources, tentatively understanding 'the good' in impartial welfarist terms" (MacAskill 2019). (Note that Shiller himself is an Effective Altruist.[18]) This leads me to believe that Ord would *at least* be indifferent to terminal extinction once posthumanity arrives. Though it seems more likely that, if pressed, he would concede that Shiller's argument goes through: we should die out after the arrival of posthumanity, since failing to do so would impede the maximization of value. My suspicion is that many longtermists hold this pro-extinctionist view: our moral-eschatological duty is to create or become posthumans who establish a techno-utopian world among the stars, described by Ord as our "vast and glorious" future (Ord 2020; Redacted). Once this duty has been discharged, there are good reasons for *Homo sapiens* to bow out.[19]

Either way, I contend that the most likely outcome of extinction neutralism would be indistinguishable from the outcome prescribed by pro-extinctionists. If we create "superior" posthumans to rule the world, what reasons would there be for *them* to keep "inferior" beings like us around? Perhaps they would place "legacy humans" in zoos or keep

us as pets, though we would, once again, continue using up resources that could be better spent on building a cosmic utopia (see Goertzel 2010; Bostrom 2020).

Alarmingly, there is almost no serious discussion among longtermists, or extinction neutralists more generally, about what life might be like for members of our species in the posthuman era. Some longtermists imagine *themselves* becoming posthuman, which renders the question less pressing from their perspective, as it won't matter to *them* what happens to legacy humans.[20] For those who decide not to become posthuman, I see no reason for expecting their lives to flourish—or to persist. There are too many reasons for posthumanity to sideline, disempower, marginalize, and ultimately eliminate our species. When the fate of *Homo sapiens* is no longer in its control, the human population will probably diminish for the same reasons that ~75% of all primate species are undergoing population decline and ~60% are now at risk of extinction (Estrada et al. 2017). The preservation of primate species is simply not a priority for humans, nor will the preservation of humans be a priority for posthumanity. This is why I believe that extinction neutralism has pro-extinctionist implications *in practice*.

Hence, it seems to me that longtermism is closely connected to pro-extinctionism: some longtermists would likely agree that value would be better maximized without our species around, especially given finite resources and the utilitarian imperative to maximize welfare in particular. But even if they were to accept extinction neutralism, the outcome would likely be the same for us: terminal extinction.

6. Conclusion

One might suspect that pro-extinctionism is a fringe view. This is true with respect to pro-extinctionist positions that endorse final human extinction. But there is an alternative form of pro-extinctionism that endorses terminal extinction without final extinction, which I have dubbed Silicon Valley pro-extinctionism. This is increasingly popular within tech circles, and hence poses a much greater threat to the continued survival of our species. A novel contribution of this paper, I hope, is how it reconceptualizes ideologies like accelerationism and longtermism as instances of, or being adjacent to, pro-extinctionism. As Adam Kirsch (2023) argues, there is a widespread "revolt against humanity," which takes many forms: Efilism, radical environmentalism, and antinatalism, as well as TESCREAL ideologies like transhumanism, of which accelerationism and longtermism are variants (Redacted).

Beyond this contribution, I hope to have provided some degree of conceptual clarity to discussions of pro-extinctionism, which I predict will become more pressing in the coming years given its rise within Silicon Valley. (Indeed, a growing number of young AI researchers now claim that it is "fundamentally unethical" to have biological children because the future is and should be digital, a view that I have elsewhere termed "replacement antinatalism"; Redacted.) Despite the wide variety of different pro-extinctionist positions, all share a common denominator: that the future ought not include our species. If one believes that our species should have a future, then one should oppose pro-extinctionism in all its forms.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Akerma, Karim. "Theodicy shading off into Anthropodicy in Milton, Twain and Kant." Tabula Rasa 41 (2010).

Beiser, Frederick C. Weltschmerz: pessimism in German philosophy, 1860-1900. Oxford University Press, 2016.

Benatar, David. Better never to have been: The harm of coming into existence. Oxford University Press, 2008.

Bostrom, Nick. “Existential risks: Analyzing human extinction scenarios and related hazards.” *Journal of Evolution and technology* 9 (2002).

Bostrom, Nick. "Astronomical waste: The opportunity cost of delayed technological development." Utilitas 15, no. 3 (2003): 308-314.

Bostrom, Nick. "Transhumanist values." Journal of philosophical research 30, no. Supplement (2005): 3-14.

Bostrom, Nick. "Letter from utopia." Studies in Ethics, Law, and Technology 2, no. 1 (2008).

Bostrom, Nick. "Existential risk prevention as global priority." Global Policy 4, no. 1 (2013): 15-31.

Bostrom, Nick. Superintelligence: Paths, Dangers, Strategies. United Kingdom: Oxford University Press, 2014.

CoE. “e-sermon #9.” Church of Euthanasia. https://www.churchofeuthanasia.org/e-sermons/sermon9.html.

Crisp, Roger. "Would extinction be so bad." The New Statesman 10 (2021): 2021.

Crisp, Roger. "Pessimism about the Future." Midwest Studies in Philosophy 46 (2023).

Estrada, Alejandro, et al. “Impending Extinction Crisis of the World’s Primates: Why Primates Matter.” *Science Advances*, 18;3(1).

Faggella, Daniel. “Potentia: The Morally Worthy ‘Stuff.’” Personal website, 2023. https://danfaggella.com/potentia/.

Faggella, Daniel. X post, 2025a. https://x.com/danfaggella/status/1963063013092512226.

Faggella, Daniel. “Taboo Cognition: Why People Hate Posthumanism.” Email, 2025b. https://ao8p8.r.a.d.sendibm1.com/mk/mr/sh/6rqJ8GoudeITQRjoZgjbRkcY59a/WqU3OZDBd9MG.

Faggella, Daniel. X post, 2025c. https://x.com/danfaggella/status/1986881298288550339.

Finneron-Burns, Elizabeth. "What’s wrong with human extinction?." Canadian Journal of Philosophy 47, no. 2-3 (2017): 327-343.

Flannery, Frances L. Understanding apocalyptic terrorism: countering the radical mindset. Routledge, 2015.

GLF. 1994. Statement of Purpose (A Modest Proposal). https://www.churchofeuthanasia.org/resources/glf/glfsop.html.

Goertzel, Ben. "A cosmist manifesto: Practical philosophy for the posthuman age." (No Title) (2010).

Greaves, Hilary, and William MacAskill. "The case for strong longtermism." Global Priorities Institute Working Paper No. 5-2021 (2021).

Harari, Yuval Noah. “Industrial Farming Is One of the Worst Crimes in History.” The Guardian, Guardian News and Media, 25 Sept. 2015, www.theguardian.com/books/2015/sep/25/industrial-farming-one-worst-crimes-history-ethical-question.

Harrison, Gerald. "Antinatalism, Asymmetry, and an Ethic of Prima Facie Duties1." South African Journal of Philosophy 31, no. 1 (2012): 94-103.

Häyry, Matti, and Amanda Sukenick. *Antinatalism, Extinction, and the End of Procreative Self-Corruption*. Cambridge University Press, 2024.

Hendrycks, Dan, Mantas Mazeika, and Thomas Woodside. "An overview of catastrophic ai risks." arXiv preprint arXiv:2306.12001 (2023).

Jezos, Beff (aka Gill Verdon) and Bayes. “Notes on E/Acc Principles,” *Substack*. https://beff.substack.com/p/notes-on-eacc-principles-and-tenets.

Khatchadourian, Raffi. “The Doomsday Invention.” *The New Yorker*, 2015. https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom.

Kirsch, Adam. *The Revolt Against Humanity: Imagining a Future Without Us*. Columbia Global Reports, 2023.

Knight, Les. These Exit Times. Voluntary Human Extinction Movement, 1991. https://www.vhemt.org/TET1.pdf.

Knutsson, Simon, and First written June. "Thoughts on Ord's 'Why I'm Not a Negative Utilitarian.'" (2016).

Knutsson, Simon. "Philosophical Pessimism: Varieties, Importance, and What to Do." APA Blog. https://blog.apaonline.org/2022/09/13/philosophical-pessimism-varieties-importance-and-what-to-do%ef%bf%bc/.

Knutsson, Simon. 2023. "My moral view: Reducing suffering, 'how to be' as fundamental to morality, no positive value, cons of grand theory, and more." https://www.simonknutsson.com/my-moral-view/.

Korda, Chris. "Credits," 1994. https://www.churchofeuthanasia.org/snuffit5/Credits.html.

Korda, Chris. "Naval Assault on Earthfest," 1994. https://www.churchofeuthanasia.org/snuffit5/Naval_Assault.html.

Kurzweil, Ray. "The singularity is near." In Ethics and emerging technologies, pp. 393-406. London: Palgrave Macmillan UK, 2005.

Ladish, Jeffrey. X post, 2025. https://x.com/JeffLadish/status/1932540354496246231.

Lovejoy, Arthur. The Great Chain of Being. Harvard University Press, 1936.

MacAskill, William. "Defining Effective Altruism," *Effective Altruism Forum*, 2019. https://forum.effectivealtruism.org/posts/9wYa8BqSTMcx9j2tK/defining-effective-altruism.

MacAskill, William. What We Owe The Future: The Sunday Times Bestseller. Simon and Schuster, 2022.

Manheim, David. X post, 2025. https://x.com/davidmanheim/status/1967205334319309062.

Marantz, Andrew. "Among the AI Doomsayers." The New Yorker, 2024. https://www.newyorker.com/magazine/2024/03/18/among-the-ai-doomsayers.

Moravec, Hans. "Human culture: a genetic takeover underway." In Artificial Life, pp. 167-199. Routledge, 2019.

Narveson, Jan. Utilitarianism and New Generations." Mind 76, no. 301 (1967): 62–72.

Narveson, Jan. "Future People and Us." In Obligations to Future Generations, edited by.

Newberry, Toby. "How many lives does the future hold?." Global Priorities Institute Technical Report (2021).

Ord, Toby. “Why I’m Not a Negative Utilitarian.” (2013). http://www.amirrorclear.net/academic/ideas/negative-utilitarianism/.

Ord, Toby. “Opening Keynote,” 2016. https://www.youtube.com/watch?v=VH2LhSod1M4.

Ord, Toby. The precipice: Existential risk and the future of humanity. Hachette Books, 2020.

Pearce, David. Hedonistic imperative. David Pearce., 1995.

Pearce, David. “The Biointelligence Explosion,” in Amnon Eden et al. (eds.) Singularity Hypotheses: A Scientific and Philosophical Assessment, 2012. Springer.

RAINN. “Victims of Sexual Violence: Statistics.” RAINN, Rape, Abuse, and Incest National Network (2024). www.rainn.org/statistics/victims-sexual-violence.

Rothman, Joshua. "The case for not being born." The New Yorker (2017).

Schopenhauer, Arthur. Essays and aphorisms. Vol. 227. Penguin UK, 1970.

Schopenhauer, Arthur. The World as Will and Representation. Aegitas, 2016.

Shiller, Derek. "In Defense of Artificial Replacement." Bioethics 31, no. 5 (2017): 393-399.

Shiller, Derek. Effective Altruism Forum profile, 2019. https://forum.effectivealtruism.org/users/derek-shiller.

Singh, Asheel. "Furthering the case for anti-natalism: Seana Shiffrin and the limits of permissible harm." South African Journal of Philosophy 31, no. 1 (2012): 104-116.

Sutton, Richard. Twitter post, 2022. https://x.com/richardssutton/status/1575619655778983936.

Sutton, Richard. “AI Succession,” 2023. https://www.youtube.com/watch?v=NgHFMolXs3U.

Tegmark, Max. Life 3.0: Being human in the age of artificial intelligence. Vintage, 2017.

Trust, Sandy, et al. “Planetary Solvency: Finding Our Balance with Nature.” University of Exeter, 2025. https://actuaries.org.uk/media/wqeftma1/planetary-solvency-finding-our-balance-with-nature.pdf.

Verdon, Gill. Twitter post, 2022. https://x.com/BasedBeffJezos/status/1599740056456941568.

Verdon, Gill. X post, 2025. https://x.com/beffjezos/status/1958681234432987327.

Vetter, Hermann. “Discussion.” In Induction, Physics, and Ethics, edited by Paul Weingartner and Gerhard Zecha. D. Reidel Publishing Company, 1968.

Vetter, Hermann. “IV. The Production of Children as a Problem of Utilitarian Ethics.” Inquiry: An Interdisciplinary Journal of Philosophy 12, no. 1–4 (1969): 445–447.

Vetter, Hermann. “Utilitarianism and New Generations.” Mind 80, no. 318 (1971): 301–2.

Walker, Mark. “Ship of Fools: Why Transhumanism Is the Best Bet to Prevent the Extinction of Civilization.” *Metanexus*, 2009. https://metanexus.net/h-ship-fools-why-transhumanism-best-bet-prevent-extinction-civilization/.

Wells, Andy. 2017. “What Is ‘Legal’ Torture and How Many Countries Still Use It?” Yahoo News. https://uk.news.yahoo.com/what-is-legal-torture-and-how-many-countries-still-use-it-165034887.html.

Yudkowsky, Eliezer. “You Only Live Twice.” *Lesswrong*, 2008. https://web.archive.org/web/20230323081901/https://www.lesswrong.com/posts/yKXKcyoBzWtECzXrE/you-only-live-twice.

[1] See section 5 for qualifications.
[2] See Redacted
[3] Put differently, in the first case, one starts with antinatalism (as an ethical position) and ends up with pro-extinctionism, whereas in the second case, one starts with pro-extinctionism and ends up with antinatalism (as a method).
[4] See footnote 14.
[5] Hence, omnicide could be either voluntary or involuntary. It is, however, almost always assumed to be an involuntary event, whereby some group unilaterally kills everyone without the consent of others.
[6] Note that Crisp, somewhat perplexingly, suggests that a massive asteroid impact would instantaneously kill everyone on Earth. Perhaps this would happen if Earth collided with a very large *planet*, but for asteroids 12 km in diameter (like the one that killed off the nonavian dinosaurs), massive human suffering would result due to the impact winter.
[7] One could theoretically argue in favor of omnicide to prevent new people from being born.
[8] Although an ethical antinatalist could prefer alternative methods to eliminate procreation, such as omnicide (rather than methodological antinatalism).
[9] Two other examples: Asheel Singh writes that, since harming someone without their consent is wrong because it would violate their rights, and since one cannot ask the unborn for consent to be born, procreation is always wrong, as everyone who is born will at some point experience some harm (Singh 2012). And Hermann Vetter notes that, on the person-affecting utilitarianism of Jan Narveson, if one were to create a child who would have a happy life, no utilitarian duty would be violated; but if one were to create a child who has a bad life, one's duty will have been violated. And since we cannot know whether our children will have good or bad lives once created, we should not create them (Vetter 1969).
[10] That said, Benatar's asymmetry is reminiscent of an asymmetry pointed out in Vetter (1971).
[11] See Häyry and Sukenick 2024.
[12] See May 2024 for further discussion.
[13] Note that I do not mention these examples — Faggella, Druggan, and (below) Verdon — because they are academically or philosophically respectable views. Rather, I consider them worth highlighting because such individuals represent views that are becoming increasingly widespread within certain powerful corners of the tech world, especially within the field of AI.
[14] For example, he writes:

> Expanding potential will uncover all that is "good." On the evolutionary journey upwards from flatworms to humans, think of all the "good" that was discovered. Creativity, love, humor, modes of communication / collaboration – so much value was uncovered as potentia expanded. Yet this is all just scratching the surface of potential value – most of the possible "goods" have not been discovered (Faggella 2023).

[15] As Faggella writes, "a safe (or at least not overtly and unnecessarily dangerous) expansion of potentia is the sole moral imperative" (Faggella 2023).
[16] For example, the Effective Altruist Jeffrey Ladish writes in a response to Faggella:

> I have a special love for humans (and other animals) and a lot of stake in the preferences of current and future humans (and other minds too). I also am pretty down for creating many other types of minds, but I have a strong preference for the existence and continuity of people alive today and their descendants (Ladish 2025).

David Manheim posted a survey on social media asking people who identify as Effective Altruists whether they endorse the replacement of humanity with posthumanity. Although not a scientific poll, it is still worth noting that 18.5% opposed AGI ever replacing humanity, 23.3% said this would be okay eventually (if AGI preserves our values), and 16.9% opted for "create whatever AGI leads to maximal utility" (Manheim 2025).
[17] A failure to fulfill this potential would result in an "existential catastrophe."
[18] See Shiller 2019.

[19] Recall from section 1 that longtermism should be classified as a "further-loss view" *if* one accepts the Broad Definition of 'humanity.' On this account, human extinction would constitute a tragedy of quite literally cosmic proportions. However, if one accepts the Narrow Definition, it may *instead* be classified as "pro-extinctionist." This is why, I argue, clearly defining terms like 'humanity' and 'extinction' is crucial in discussions about the ethics of human extinction (Redacted). Since the present paper has assumed the Narrow Definition, longtermism could indeed be a form of pro-extinctionism—though, as noted, it is also compatible with extinction neutralism and the coexistence view.

[20] This is indicated by the fact that individuals like Nick Bostrom, Eliezer Yudkowsky, and others in the community have signed up with cryonics companies to have their brains cryogenically frozen so they can be resurrected in the future (Khatchadourian 2015; Yudkowsky 2008;).