

Whose Survival? A Critical Engagement with the Notion of Existential Risk*

PHILIP HØJME

Polish Academy of Sciences, IFIS, GSSR, Poland

contact.hoejme@gmail.com

ORCID: 0000-0003-2861-9027

Abstract. This paper provides a critique of Bostrom's concern with existential risks, a critique which relies on Adorno and Horkheimer's interpretation of the Enlightenment. Their interpretation is used to elicit the inner contradictions of transhumanist thought and to show the invalid premises on which it is based. By first outlining Bostrom's position this paper argues that transhumanism reverts to myth in its attempt to surpass the human condition. Bostrom's argument is based on three pillars, Maxipok, Parfitian population ethics and a universal notion of general human values. By attempting to transcend the human condition, to achieve post-humanity, transhumanism reverts to myth. Thus, the aim of this paper is to provide a critical examination of transhumanism which elicits its tacit contradictions. It will also be argued that transhumanism's focus on a universal, all-encompassing, notion of humanity (Earth-originating intelligent life) neglects any concern with actual lived lives. This absence is problematic because it clearly shows that there is a discrepancy, between transhumanism's claimed concern for all of humanity and the practical implications of proposing a universal notion of

* Certain ideas expressed in this paper has previously been presented at conferences, as indicated in the text. The critique of transhumanism based on Butler's Adorno Prize Lecture were originally conceived as part of a tutorial on existential risks held at the University of Amsterdam, The Netherlands, supervised by Stefan Niklas. Additionally, I would like to thank the research group supervised by Adam Lipszyc, as well as my friend Jakob Stenske, for their critical comments.

humanity. This paper will conclude, that transhumanism's lack of concern with actual lives is due to its universal and totalising gestures. Gestures which allow for universal claims such as general values or Earth-originating intelligent life.

Keywords: critical theory; enlightenment; future; humanity; transhumanism.

The threats posed to humanity has recently shifted from concerns with nuclear annihilation (which was the concern during the Cold War) to a concern with the threats posed to humanity by e.g. climate change, AI, asteroids colliding with Earth, and other perceived global threats. It is with these concerns in mind that this article seeks to critique the Transhumanist notion of the importance of saving humanity. The structure of the article is as follows: firstly, a brief introduction to the development of Transhumanism – its origins and current developments, secondly, an examination of Nick Bostrom's claim that the survival of the human species is of paramount importance for humanity – humanity ought to be concerned with existential risks. This position will then be critiqued using both, Theodor W. Adorno and Max Horkheimer's critique of the myth of enlightenment, and by utilising Judith Butler's notion of grievability against the universal notion of humanity which is an integral part of the Transhumanist ideology. The main gist of the critique levelled against the Transhumanist project is that it lacks a concern with actual lives (as opposed to abstract lives) and that it in its attempt to transcend the human condition reverts to the myth of the grandeur of humanity's possibilities, the myth of progress.

In *A History of Transhumanist Thought* (2005a), Bostrom provides a genealogy of transhumanism, arguing that the idea of extending one's life can be traced back to the *Epic of Gilgamesh* (approx. 1700 B.C.). While stating that modern transhumanism began with the publication of Darwin's *Origin of Species* (1859) and Nietzsche's conception of the Übermensch, in *Thus Spoke Zarathustra* (1883). Both books were pivotal in suggesting an understanding of humanity which does not see the current humanity as "the endpoint of evolution but rather as a possibly quite early phase" (Bostrom, 2005a, 3). An early phase meaning that it is but a stepping stone towards a more conscious

development (as is seen in the growing numbers of biohackers) or the natural evolution of humanity (Bostrom, 2003). In the years following the second world war, science-fiction seems to have caught the imagination of many, and writers such as Karel Čapek, Isaac Asimov, Stanisław Lem, and Arthur C. Clark (Bostrom, 2005a, 7) became important figures in shaping many people's conception of the future. The book *Are you a transhuman?* (1989) written by FM-2030 (formerly F.M. Esfandiary) and the 1998 founding of The World Transhumanist Association, by Bostrom and others, constitutes two decisive moments in transhumanist history. The founding of the World Transhumanist Association culminated in the publication of the transhumanist declaration (HumanitiesPlus, 2009). A declaration of great importance for setting out the goals and responsibilities of transhumanism. Bostrom (2005a) also goes to great lengths to elaborate on the Enlightenment roots of transhumanism. Clearly emphasising both transhumanism's concern with individual liberties and stressing its liberal utilitarian concern for the welfare of all humans. A concern which is found in J.S. Mill's *Utilitarianism* (1863) rather than in the writings of Nietzsche. Such emphasis seems to suggest two points, that transhumanism is opposed to eugenics, and that it does not condone a totalising state ideology (e.g. fascism or nazism). Such a concern surface when Bostrom writes that "The Holocaust left a scar in the human psyche" (Bostrom, 2005a, 6). The Holocaust is understood as the extreme consequence of eugenics, as perpetrated by the nazis during the second world war. This point seems parallel to Adorno's question, posited in *Negative Dialectics* (1966), *how is one to live on after Auschwitz?* Bostrom understands the enlightenment roots of transhumanism in the broadest possible sense, as scientific and cultural progress, and sees the explicit aim of this logic to be a transcendence of the human condition. Hence transhumanism is a continuation of the Enlightenment project because it seeks to develop humanity using science and innovations to better the human condition. An understanding of the Enlightenment which is contrasted with Adorno and Horkheimer's interpretation. This critique, together with a critique based on the general argument in Butler's Adorno Prize Lectures (2012), forms the critical part of this paper. Adorno and Horkheimer's

interpretation of the Enlightenment can be described as claiming that myth and Enlightenment constitutes a dialectical interplay of forces. In their book *Dialectic of Enlightenment* (1944) the authors argue that rationality's overcoming of myth is, in fact, a return of the mythical, because it either subsumes or discards the non-rational by its totalising gestures. These gestures reorganize the mythical as the hidden truth in the secularized notion of progress that underlines Enlightenment thought. Opposed to such universalism Butler's argument will discern an absence in transhumanism, a lack of concern with particularity, present lives. A lack which shows that the goal of transhumanism, to save humanity, is an all-subsuming end, i.e. a myth. An end which excludes, or delegates, less priority to all other ends. While it might be said that achieving post-humanity is nonproblematic because it is a rational wish or the natural telos of evolution, this is not the only conception of post-humanity. In *The Posthuman* (2013) R. Braidotti argues for a radically different conception of the post-human. A conception which does not understand the post-human as something towards which humanity (can) strive. Instead, it is a "qualitative shift" (Braidotti, 2013, 2) of what is perceived as being the overarching commonality of the species. As such this interpretation differs greatly from the transhumanist perspective (see Bostrom 2003). And while a discussion of these two radically different conceptions of post-humanity is both interesting and important, it is outside of the scope of this paper to provide such a discussion.

Bostrom's concern depends on three arguments: a typology of risks, which leads to the realization that existential risks pose the greatest risk to humanity, a maxim to maximize the chance of an okay outcome, and what I have previously termed Parfitian population ethics (Højme, 2018. Højme, 2019). All these arguments can be found in Bostrom's paper *Existential Risk Prevention as Global Priority* (2013), but also in earlier papers such as *The Future of Humanity* (2009) and *Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards* (2002).

The typology of risk offered is structured along two axes; intensity and scope, totalling six categories of risks (Bostrom, 2002, 0–2). Along the intensity axis, these are classified as either personal, local or global. While

along the other axis each category is conceived of as being either endurable or terminal. As examples of these six kinds of risks Bostrom gives the following examples: having one's car stolen (personal-endurable), a fatal car crash (personal-terminal), recession in a country (local-endurable), genocide (local-terminal), thinning of the ozone layer (global-endurable) and X (global-terminal) (Bostrom, 2002, 1). Bostrom later expands the number of categories from six to fifteen, renaming the intensity axis scope, and scope axis severity (Bostrom, 2013, 17). This is done without any noteworthy change to the general argument of what constitutes existential risks. Instead, the development simply constitutes a clarification of the typology into further subdivisions. Existential risks are the sixth/fifteenth kind, classified under the header 'X' in both papers. These risks differ from the global-endurable kind because they produce "an adverse outcome [which] would either annihilate Earth-originating intelligent life or permanently and drastically curtail its potential" (Bostrom, 2002, 2). Asteroids and comets, among other events – e.g. climate change, AI or nuclear war – could cause such an extinction. While the threat of these risks is not neglected, Bostrom argues that it is in fact technology which currently poses the biggest threat to humanity. "The first manmade existential risk was the inaugural detonation of an atomic bomb" (Ibid.), since which the threats posed to humanity by technology have only increased. The main thrust of Bostrom's argument seems to be that there is a need for proactive action in the case of existential risks, because "Our approach to existential risks cannot be one of trial-and-error." (Ibid. 3). In the case of existential risks, humanity cannot learn from its errors, because such errors foreclose, or annihilates, the future. Additionally, existential risks are international, and as such we cannot let national policies dictate our reaction to these risks. This means that "Reductions in existential risks are *global public goods*" (Ibid. 4), and the reduction of these risks would benefit all of humanity. Saving current, and future, generations from extinction by reducing the likelihood of existential risks amounts to creating a utility surplus much greater than would, for example, saving the lives of 10,000 people a year by inventing a new airbag. This is what was meant by Parfitian population ethics. When claiming that

a fatal car crash is less severe than the extinction of humanity, it proposes that humanity ought to be more concerned with existential risks than car crashes. The rationale behind this argument is taken directly from Derek Parfit's claim that: an event which exterminates humanity is worse than an event that 'only' annihilates 99.9% of the population (Bostrom, 2013, 17–18, Parfit, 1984, 453). If Parfit is correct in this assumption, and Bostrom clearly thinks this is the case; humanity ought to avoid extinction at all costs since extinction is infinitely worse than the loss of 99.9...% of humanity. This claim thus forms an important premise for Bostrom's claim that humanity ought to prioritise a reduction in the likelihood of existential risks.

Further expanding on the notion of existential risks, Bostrom offers a taxonomy of the different categories of existential risks. Postulating four general categories, each with numerous secondary kinds, a description of which is omitted because the main concern of this paper is a detailed, but not an extensive outline of existential risks.

Bangs – Earth-originating intelligent life goes extinct in relatively sudden disaster resulting from either an accident or a deliberate act of destruction.

Crunches – The potential of humankind to develop into posthumanity is permanently thwarted although human life continues in some form.

Shrieks – Some form of posthumanity is attained but it is an extremely narrow band of what is possible and desirable.

Whimpers – A posthuman civilization arises but evolves in a direction that leads gradually but irrevocably to either the complete disappearance of the things we value or to a state where those things are realized to only a minuscule degree of what could have been achieved (Bostrom, 2002, 5).

So far, the primary example of existential risks has been bangs. But each kind of existential risks constitutes a calamity for humanity because they either curtail the transhumanist agenda, of achieving posthumanism, or wipe out humanity before this condition is reached. Such prospects led Bostrom to propose a maxim to guide our prioritization when choosing which risks should be awarded most attention. A maxim which depends

on Parfitian population ethics as previously described. Bostrom names this maxim *Maxipok* and proposes that humanity adopt a stance towards existential risks of always trying to “Maximise the probability of an ‘OK outcome’, where an OK outcome is any outcome that avoids existential catastrophe” (Bostrom, 2013, 19). This does not mean that *Maxipok* is an imperative maxim; instead, it is a rule of thumb. This seems to be a tactical move because it is treated elsewhere as an imperative:

usefulness [as a way of deciding between options] is as an aid to prioritisation. [Because] Unrestricted altruism is not so common that we can afford to fritter it away on a plethora of feel-good projects of suboptimal efficacy. If benefiting humanity by increasing existential safety achieves expected good on a scale many orders of magnitude greater than that of alternative contributions, we would do well to focus on this most efficient philanthropy (Ibid. 19).

This statement does, provoke one to question *whether the argument is not construed to make sure that Maxipok always comes out being the most reasonable of choices?* Chiefly because *Maxipok*’s underlining validity is based on Parfitian population ethics, which in turn seems to play a (arbitrary) numbers game which can only lead to the conclusion Bostrom wishes for. Thus, Bostrom’s claim that it is a rule of thumb seems more like a ruse against criticism of the rule’s inherent primacy given to the mitigation of existential risks. A primacy which facilitates a lack of interest in the lives currently lived.

So far in this paper, the usage of humanity and Earth-originating intelligent life has been synonymous, but it is important if one is to understand Bostrom’s concern that these two concepts are differentiated. While humanity is conceived of, in the classic sense of the word, as *homo sapiens*, Bostrom’s argument is rather concerned with Earth-originating intelligent life, because “there is no reason to suppose that the biological species concept tracks what we have reason to value” (Ibid. 20). This paper will not engage in the discussion of the classification of current humanity, even if others (including transhumanists) are engaged in these discussions (see e.g. the notion of *homo sapiens technologicus* [Kermisch, 2011. *Delicata* 2018.]).

To encompass such diverse views of humanity, and what it might become in the future, Bostrom proposes using the term Earth-originating intelligent life (a term which includes future possible developments/evolution of humanity), as opposed to the more limited notion of humanity (as it is used colloquially). Such a proposition is open to the following criticism: *What are the values which could be shared by future kinds of "humanity"?* Bostrom suggests a broadening of the concern for humanity which also includes future conditions. Conditions which might be unimaginable in the present or noncompatible with our current views on what constitutes humanity. By framing the argument in this way Bostrom seems to safeguard against critics who either do not believe there is such a notion as humanity or critics who (like the transhumanists themselves) see the purpose of humanity as being a transcendence of this condition.

Taking up the discussion of values in the book *Superintelligence* (Bostrom, 2014a), Bostrom argues that it might be important for a superintelligence to learn the values of humanity. While it is possible to create a program with certain capabilities (see capabilities approach, Bostrom, 2014a, 2014b), meaning that the superintelligence comes programmed with allowed or barred actions for specific scenarios, such a comprehensive program would be an impossible task to code (Bostrom, 2014a, 282). Thus, it is a far better solution to teach a superintelligence the general values of humanity (for a thorough discussion of the problem of value-loading see Bostrom, 2014b). There are, however, problems with teaching superintelligence the general values of humanity. One of the chief problems would be that humans hold diverse and sometimes immensely complex values (even if they seem simple to us, these values might be very complex to describe in code). Another problem is that of translating ordinary language into code (Bostrom, 2014a, 283). Both problems suggest that "we cannot transfer human values into an AI by typing ... representations in computer code" (Ibid. 284). By analysing different schemes for solving these problems Bostrom argues that some schemes are potentially useful, while others are conceived as being ill-suited for this purpose. While this discussion is interesting, our present concern limits us to Bostrom's attempt at finding

a criterion for choosing values (Ibid. 209–227). Here Bostrom posits that there is “No ethical theory [which] commands majority support among philosophers, so most philosophers must be wrong [regarding ethical issues] ... [Thus, it is] Very likely, [that] we are still labouring under one or more grave moral misconceptions” (Ibid. 318). As a solution to this problem, Bostrom suggests that humanity takes up “indirect normativity ... [because] we may not know what we truly want ... [thus,] we would delegate some of the cognitive work required for value selection to the superintelligence” (Ibid. 319). A solution ‘similar’ to Bostrom’s seems to have been ‘proposed’ by Douglas Adams in *The Hitchhiker’s Guide to the Galaxy* (Adams, 2003). Here a supercomputer is tasked with finding the answer to the somewhat underdetermined question ‘of life, the universe and everything’. However, this does seem to come with its own issue of, *how will the computer know what the question is?* In Adams’ book, the answer ‘forty-two’ does not satisfy those who asked the question (philosophers). Hence, this issue suggests that even a superintelligence might not be able to help us find the answer to the ultimate question, let alone to the question of which values should have priority for Earth-originating intelligent life. This is a problem that Bostrom is aware of, stipulating that it might be advantageous to review any suggestions of the superintelligence. “The main purpose of ratification would be to reduce the probability of catastrophic error. In general, it seems wise to aim at minimizing the risk for catastrophic error” (Bostrom, 2014a, 345). But if human revision is needed, then *how does this solution solve the problem of diverse interpretations of which values are to be prioritised?* Moreover, this would make the superintelligence subordinate to the fallibility of human knowledge. Once again, we arrive back at the original question: *Which values are the most important for humanity? And how do we determine this?* What has been shown so far, is that the lack of a precise understanding of which values are to take priority, coupled with the arbitrariness of Parfitian population ethics, makes the concerns of transhumanism rather normative. Normative because the argument for the priority of existential risks suggests that there are universal values. Values which overrides other, less important concerns. This means that transhumanism’s aim to ‘cheat’ death, to avoid extinction

of Earth-originating intelligent life, is conceived of as the essential telos of humanity/Earth-originating intelligent life. As a critique of this position, I will offer an alternative interpretation of the Enlightenment, posited by Adorno and Horkheimer. In addition to this, a critique of the lack of concern with the actual lived lives is offered, based on Butler's conception of greivablity, this critique revolves around the question: *Whose survival is it Bostrom is concerned with?*

Adorno and Horkheimer's interpretation of the Enlightenment was an attempt to dispel metaphysical truth claims. They demonstrated how the Enlightenment's demystification of *a priori* categories, reverted to myth by stipulating new universal categories – "human rights then fare no better than the older universals ... Enlightenment is totalitarian" (Adorno & Horkheimer, 2002, 3–4). Hence, the Enlightenment became as totalitarian as the myths it dispelled, by relying on "identification ... thinking that tolerates nothing outside it" (Adorno, 1990, 172). For Adorno and Horkheimer the human endeavour to live without fear becomes a wish for absolute knowledge, it attempts to unify everything within the realm of what is known. Leaving nothing outside of its epistemological sphere, except non-existence, the Enlightenment seeks an "immanence of positivism ... since the mere idea of the 'outside' is the real source of fear" (Adorno & Horkheimer, 2002, 11). The Enlightenment was poised to revert to myth the moment it attempted to break the mythical spell. By postulating all to be knowledgeable, rationalization made supreme knowledge of mathematical knowledge, which in turn became a (new) myth to suppress humanity. "[The] enlightenment believes itself safe from the return of the mythical. It equates thought with mathematics. The latter is thereby cut loose, as it were, turned into an absolute authority" (Ibid. 18). Understanding the Enlightenment in this way, and it is my suggestion that we should, Bostrom's notion of universal values is a case of the Enlightenment reverting to the mythical. Parfitian population ethics is an instance of absolute mathematical truth par excellence. It is this rationalized mathematical truth that provides the premise for the priority given to existential risks. An idealising which makes existential risks the *prima facie* (absolute) concern of humanity. If Parfitian population ethics

are called into question, the primacy of existential risks loses its claim to universal validity. Additionally, by unifying humanity under the broader header of Earth-originating intelligent life, transhumanism's concern with salvation becomes another 'act' of subjection because "the oppression of society always bears the features of oppression by a collective" (Ibid. 16). Hence, the claim that transhumanism has its roots in the Enlightenment comes to mean that transhumanism's idea of the progress, or transcendence of, the human condition is built on a myth, the myth of a unified category of humanity which can be transcended. Criticising such universal conceptions Adorno stipulates that: "objects do not go into their concepts without leaving a remainder ... it indicates the untruth of identity" (Adorno, 1990, 5). This means that universal concepts – human rights, Earth-originating intelligent life, and others – elicit their own untruth. The remainder left outside suggests another critique of Bostrom's project, a critique which can be gleaned from Butler's Adorno Prize Lecture: *Can One Lead a Good Life in a Bad Life?* (Butler, 2012).

Butler's main thrust is an inquiry into what demarcates lives as being worth living. Butler suggests that grievability is paramount in such an inquiry, defining grievability as: "[If] I have no certainty that I will have food or shelter, or that no social network or institution would catch me if I fall, then I come to belong to the un-grievable" (Ibid. 10). Hence, a life worth living is a life that would constitute a loss if lost. This is, however, not the case of all lives, since "One can survive without being able to live one's life. And in some cases, it surely does not seem worth it to survive under such conditions" (Ibid. 15). This suggests two critiques of Bostrom's concern with existential risks. First, it suggests that there is an issue of who will survive. Second, it shows the agenda's complete disassociation with a concern for un-grievable lives. Bostrom attempts to mitigate such criticism by claiming that if "someone is excluded from the original extrapolation base [this] does not mean that their wishes and well-being are disregarded ... Nevertheless, it is possible that the interests of those who are included ... would be accommodated to a greater degree than the interest of outsiders" (Bostrom, 2014a, 328). Claiming that exclusion does not entail disregard

for wishes and well-being does not pass muster. Affording Bostrom the benefit of the doubt, it might be the case for those who are not included, but the excluded are never subsumed under the universal, so their wishes and desires can hardly be counted as acknowledged. More directly connected with the overall topic of transhumanism, we find Adorno's comment at the end of *Negative Dialectics* which states that: "Theologians have been unable to refrain from childishly pondering the consequences of rocket trips for their Christology, and the other way round, the infantile interest in space travel brings to light the infantilism that is latent in messages of salvation" (Adorno, 1990, 399). This infantilism, in Adorno's words, can be paraphrased as an attempt to transcend, move beyond, or surpass the reality of lived life. It is a mere distraction from the facts of late capitalism which cannot fulfil what it promised: salvation from nature, death and calamity.

The critical points raised in this paper are based on certain shortcomings inherent to Bostrom's propositions: the universal conception of Earth-originating intelligent life, the claim of universal values of humanity, and a lack of concern with those lives (currently) deemed unworthy of living. For the sake of argument, let's imagine that an asteroid was about to hit the Earth in six months and that the only option for humanity to survive would be to send spacecrafts from Earth to start a colony elsewhere. *Would it not be overly optimistic to think that an attempt at selection would be concerned with saving the most diverse representations of humanity? Would it not rather be more realistic to suppose that the majority of those aboard would be those wealthy, powerful or connected enough to secure access to, or buy, seats on the limited spacecrafts available?* Such a claim does not seem far-fetched as governments during the Cold War had contingency plans prioritising the survival of VIPs, royalty, government officials and the top brass, in the eventuality of limited or all-out nuclear war (see e.g. Vice, 2017). Transhumanism might seem like a step towards a utopian dream, of securing the preservation and continuation of Earth-originating intelligent life. As was the case with the Enlightenment, transhumanism, by naturalising humanity's linear progression towards a post-human condition, reverts to a secularised version of a universal telos of humanity. Such a vision is not

utopian, it is mythical. Instead “Utopia would be above identity and above contradiction; it would be a togetherness of diversity” (Adorno, 1990, 150). Hence, if transhumanism really seeks a utopian future, this is not to be found in a transcendence of the human condition. Instead, it is to be found where the object comes to be non-identical with its concept. This seems to offer a suggestion for the transhumanist endeavour; namely, that its critical potential should be sought where transhumanism shows the inner untruth of the current conception of humanity. Such a reorganisation would be relieved of totalising gestures, and thus it might provide a glimpse of post-humanity which would not posit any exact visions of the future. It would provide glimmers of what the future could be, flickers of utopian maybes.

Bibliography

- Adams, Douglas. 2003[1985]. “The Hitchhiker’s Guide to the Galaxy: The Original Radio Scripts”. ed. Geoffery Perkins. London: Pan Books.
- Adorno, W, Theodor. 1990[1966]. ”Negative Dialectics”. trans. E.B. Ashton. London: Routledge.
- Adorno, W, Theodor & Max Horkheimer. 2002[1944]. “Dialectic of Enlightenment”. Ed. G. S. Noerr. trans. E. Jephott. Stanford: Stanford University Press.
- Bostrom, Nick. 2014a. “Superintelligence.” Oxford: Oxford University Press.
- Bostrom, Nick. 2014b. ”Hail Mary, Value Porosity, and Utility Diversification”. Accessed March 14, 2019. <https://nickbostrom.com/papers/porosity.pdf>
- Bostrom, Nick. 2013. “Existential Risk Prevention as Global Priority.” *Global Policy*, 4(1):15–31.
- Bostrom, Nick. 2009. “The Future of Humanity.” In *New Waves in Philosophy of Technology*, ed. Jan-Kyrre Berg Olsen, Evan Selinger, & Søren Riis. 186–215. New York: Palgrave MacMillan.
- Bostrom, Nick. 2005a. “A History of Transhumanist Thought”. *Journal of Evolution and Technology*. 14(1).
- Bostrom, Nick. 2003. “Transhumanist Values”. In *Ethical Issues for the Twenty-first Century*, ed. Frederick Adams. 3–14. Charlottesville: Philosophy Documentation Center Press.
- Bostrom, Nick. 2002. “Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards”. *Journal of Evolution and Technology*, 9(1).
- Braidotti, Rosi. 2013. “The Posthuman”. Cambridge: Polity Press.

- Butler, Judith. 2012. "Can one lead a good life in a bad life? – Adorno Prize Lecture". *Radical Philosophy*, 176. pp. 9–18.
- Delicata, Nadia. 2018. "Homo technologicus and the Recovery of a Universal Ethic: Maximus the Confessor and Romano Guardini". In *Scientia et Fides vol 6, no 5*. pp. 33–53.
- HumanityPlus. 2009. "Transhumanist Declaration". Accessed March 5, 2019. <https://humanityplus.org/philosophy/transhumanist-declaration/>
- Højme, Philip. 2019. "Whose survival? A Critical Engagement with Transhumanism" Presented at Transhumanizm – Idee / Strategie / Wątpliwości 2.0, Warsaw, Poland, January, 19–20.
- Højme, Philip. 2018. "Whose survival? An exploration of the nature/culture distinction in transhumanist thought concerned with the future of humanity" Presented at The 10th Beyond Humanism Conference, Wrocław, Poland, July 18–21.
- Joseph Bien-Kahn. 2017. "You'll Die in a Nuclear War, but These Elites Will Be Saved". *Vice*, May 3. Accessed March 11, 2019. https://www.vice.com/en_us/article/wnwq4y/youll-die-in-a-nuclear-war-but-these-elites-will-be-saved
- Kermisch, Céline. 2011. "Homo Sapiens Technologicus". Review of *Homo Sapiens Technologicus*. *Techné*, 15(3):241–242.
- Parfit, Derek. revised ed. 1984[1984]. "Reasons and Persons". Oxford: Clarendon Press.