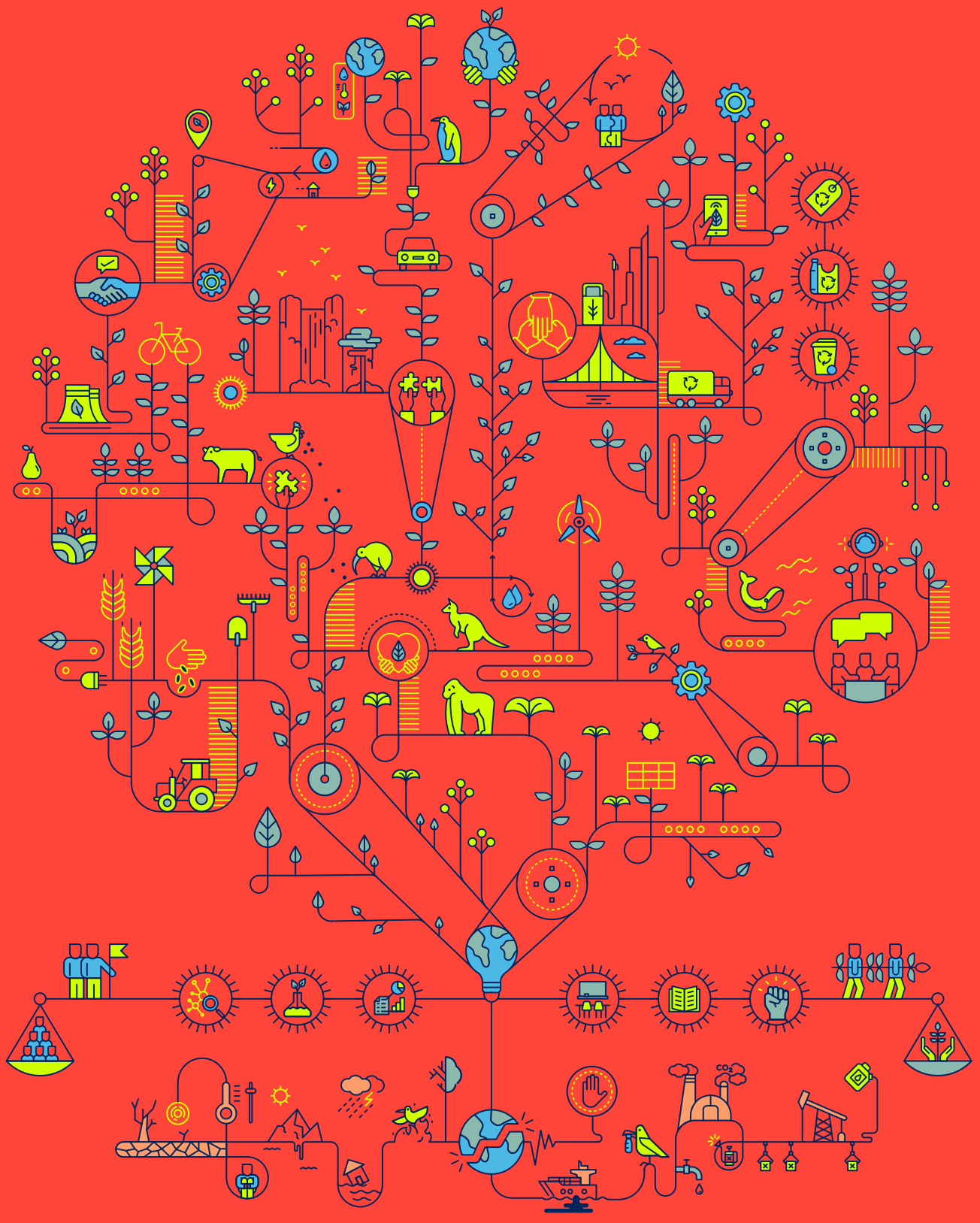
Human Development Report 2020

The next frontier

Human development and the Anthropocene



SPOTLIGHT 1.3

Existential risks to humanity

Toby Ord, Senior Research Fellow, The Future of Humanity Institute, University of Oxford

Humanity has a vast history, spanning hundreds of thousands of years. If all goes well, we can look forward to a future of equal or greater length. And just as our past saw profound expansions in our capabilities—through our lifespans, our education, our prosperity and our freedoms—so the future offers the possibility for this development to continue. We have the potential for every place on Earth to reach the highest standards seen today and to continue far beyond what has yet been achieved.

But this potential is at risk. Like every species, humanity has always been subject to the risk of extinction from natural catastrophes. And to this we have added risks of our own. Humanity's power over the world around us has increased tremendously over the past 200,000 years. In the 20th century, with the development of nuclear weapons, we became so powerful that we posed a threat to our own continued survival. This risk declined with the end of the Cold War but did not disappear. And it was joined by other risks that could threaten our continued existence, such as extreme climate change.

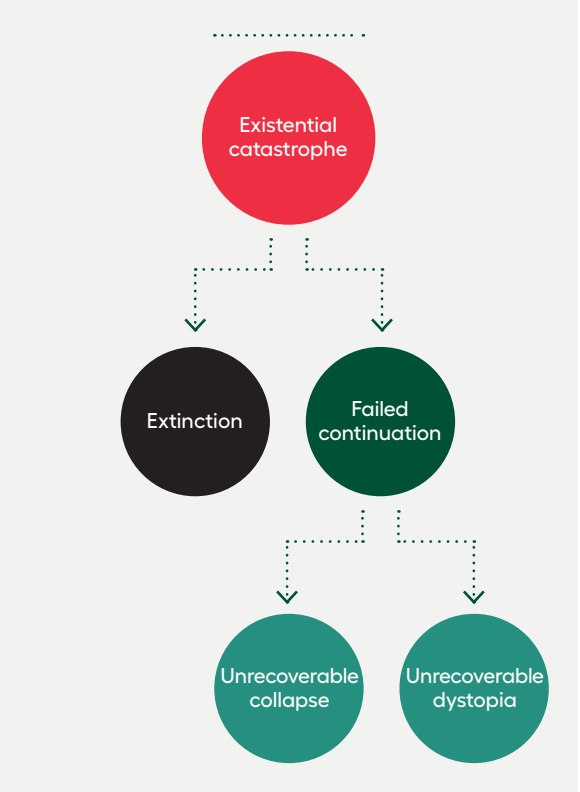
The 20th century thus ushered in a new period in which humanity has acquired the power to end its story without yet achieving the collective wisdom to ensure it does not. This period of heightened risk, known as the Precipice,[1] is closely related to the Anthropocene—indeed one suggested definition for the Anthropocene would have them begin at the same moment: 16 July 1945, when the first atomic bomb was detonated. Just as the Earth has entered a geological period in which humanity is the dominant force shaping the planet, so humanity has entered a historical period in which the dominant risks to its survival come from humanity itself. Both periods were triggered by our increasing power but may end at very different times: We could imagine a future in which humanity has found a path to safety, creating new institutions to govern global risks, such that while humanity continues to shape the planet, it has ceased to pose a substantial risk to itself.

To understand humanity's predicament, it is helpful to define two terms:

- An existential catastrophe is the destruction of humanity's long-term potential.
- An existential risk is a risk that threatens the destruction of humanity's long-term potential.[2]

The most obvious form of existential catastrophe would be human extinction, for it is clear how that would permanently foreclose our potential (figure S1.3.1). But there could be other forms too. A global collapse of civilization would also count, if it were so deep and unrecoverable that it destroyed

Figure S1.3.1 Three types of existential catastrophe



Source: Reproduced from Ord (2020).

(most of) humanity's potential. And it may also be possible for civilization to survive but be drawn into an unrecoverable dystopian future, with little value remaining.

What these outcomes have in common is that they would foreclose the possibility of human development. If such a catastrophe occurred even once, the great gains we have achieved would be permanently undone, and the possibility of reaching a more equal or more just world would be gone forever. Such risks thus threaten the most basic foundations on which almost all other value rests.

The risks

What risks could pose such a threat to our long-term potential? The most well understood are the natural risks. Take the possibility of a large asteroid impact. The mass extinction at the end of the Cretaceous 65 million years ago is widely agreed to have been caused by an asteroid, 10 kilometres in diameter, colliding with the Earth. The impact threw vast amounts of dust and ash into the stratosphere—so high that it could not be rained out. Atmospheric circulation spread this dark cloud around the planet and caused a massive global cooling, lasting years. The effects were so severe that all land-based vertebrates weighing more than 5 kilograms were killed.[3]

Scientists now have a good understanding of the chance that such an asteroid could hit us again. It is reassuringly low (table S1.3.1). In a typical century the chance of being struck by a 10 kilometre across asteroid would be just 1 in 1.5 million.[4] What about the next 100 years in particular? Scientists have modelled the orbits of all four known near-Earth asteroids of that size and confirmed that they will not hit the Earth in the next 100 years. So the remaining chance lies in the unlikely possibility that one remains undiscovered. The situation is somewhat less reassuring with asteroids between 1 and 10 kilometres across, for which detection and tracking are incomplete. Fortunately, they would also be less likely to cause a truly unrecoverable catastrophe.

Asteroids are the best-understood existential risk. They clearly pose a risk of human extinction (or unrecoverable collapse), but the risk is well understood and small. Moreover, they are the best managed existential risk: There is an effective international research programme directly working on detecting and understanding these threats.

Table S1.3.1 Progress in tracking large near-Earth asteroids

Asteroid diameter	Number	Percentage found	Chance of being struck in an average century	Change of being struck in next century
1–10 kilometres	~920	~95	1 in 6,000	1 in 120,000
10 or more kilometres	~4	> 99	1 in 1.5 million	< 1 in 150 million

Source: Adapted from Ord (2020).

There are several other known natural existential risks, including comets and supervolcanic eruptions. These are less well understood than asteroids and may pose a greater risk. Because most of these risks were discovered only within the last century, there are presumably unknown natural risks too.

Fortunately, there is a way of using the fossil record to estimate an upper bound for the total extinction risk from all natural hazards—including those that have not yet been discovered. Since humanity has survived the entire array of natural risks for thousands of centuries, the chance of extinction per century must be correspondingly small. This produces a range of estimates depending on how broad we take "humanity" to be (table S1.3.2). We can also estimate this natural extinction risk via how long related species have survived, with a range of estimates depending on how closely related they are (table S1.3.3). Both techniques suggest that the total natural extinction risk is almost certainly below 1 in 300 per century and more likely to be 1 in 2,000 or lower.[5]

Unfortunately, there is no similar argument to help estimate the total anthropogenic risk because the track record is too short. Surviving 75 years since the invention of nuclear weapons does very little to

Table S1.3.2 Estimates and bounds of total natural extinction risk per century based on how long humanity has survived, using three conceptions of humanity

Conception of humanity	Years	Best guess of risk	99.9 percent confidence bound
Homo sapiens	200,000	< 1 in 2,000	< 1 in 300
Neanderthal split	500,000	< 1 in 5,000	< 1 in 700
Homo	2,000,000 – 3,000,000	< 1 in 20,000	< 1 in 4,000

Source: Adapted from Ord (2020).

Table S1.3.3 Estimates of total natural extinction risk per century based on the survival time of related species

Species	Years	Best guess of risk
Homo neanderthalensis	200,000	1 in 2,000
Homo heidelbergensis	400,000	1 in 4,000
Homo habilis	600,000	1 in 6,000
Homo erectus	1,700,000	1 in 17,000
Mammals	1,000,000	1 in 10,000
All species	1,000,000–10,000,000	1 in 100,000–1 in 10,000

Source: Adapted from Ord (2020).

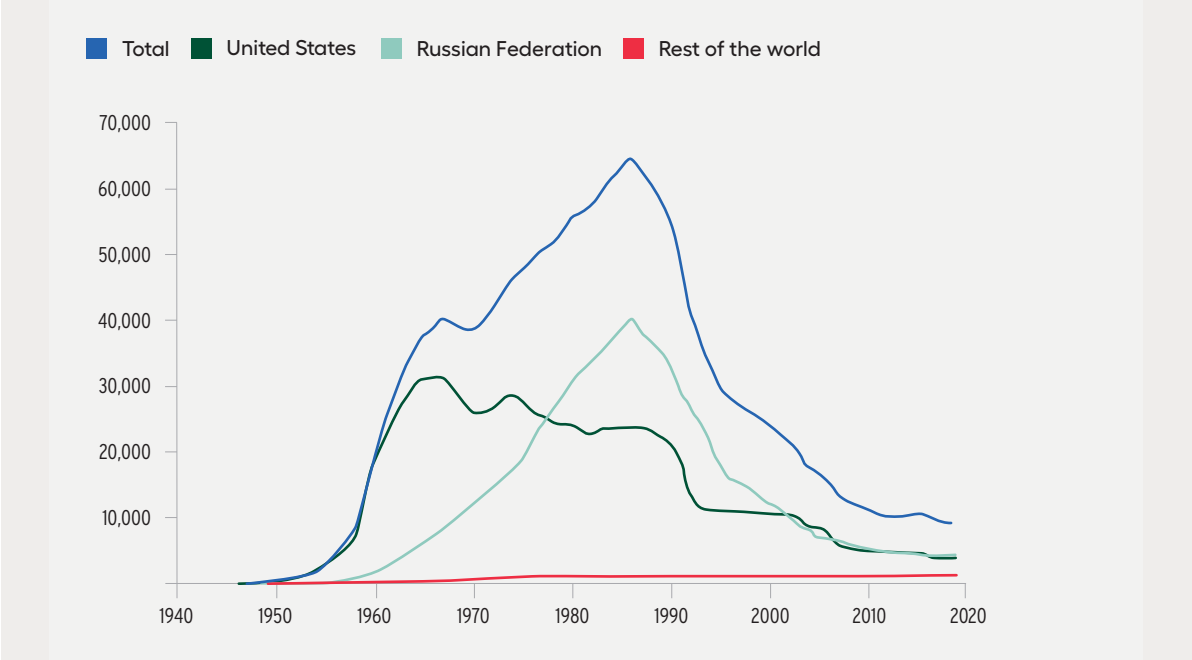
constrain the amount of existential risk from nuclear weapons over a century. We therefore have to confront the possibility that this risk may be substantial.

In the early 1980s scientists discovered that nuclear war could create a global cooling effect similar to that of large asteroid impacts.[6] While initially controversial, subsequent research has mostly supported this "nuclear winter" effect in which ash from burning cities would rise into the stratosphere, causing severe cooling lasting for years.[7] This would cause massive crop failures and widespread starvation. Researchers studying nuclear winter now suggest that a collapse of civilization might be possible, though it would be very difficult for nuclear winter to directly cause human extinction.[8]

Fortunately, the existential risk posed by nuclear war has been declining. Since the late 1980s the size of the nuclear arsenals has been substantially reduced, lowering the severity of an ensuing nuclear winter (figure S1.3.2). This appears to stem in part from concern about the existential risk the weapons posed, with both US President Ronald Reagan and USSR General Secretary Mikhail Gorbachev reporting that the possibility of nuclear winter weighed heavily on their minds.[9] Another major reduction in risk was the end of the Cold War, which has reduced the chance that the arsenals will be used at all. However, the chance has by no means been eliminated: Nuclear war could still begin through an accidental launch (and retaliation) or if tensions between great powers flare up once more.

Climate change may also pose an existential risk to humanity. Much of the scientific focus has been on the most likely scenarios. While these could be devastating by any normal measure, they would not be existential catastrophes. But some of the extreme possibilities may reach that threshold. For example,

Figure S1.3.2 While there have been substantial reductions in the number of active stockpiled nuclear warheads, the total number—especially in the Russian Federation and the United States—remains high



Source: Reproduced from Ord (2020) and adapted from Kristensen and Korda (2019).

we cannot yet rule out climate feedbacks taking us substantially beyond 6 degrees Celsius of warming—perhaps as far as 10 degrees Celsius or more.[10] It would be extremely valuable to have a better idea of the likelihood of such extreme scenarios and of whether civilization, or humanity itself, would survive them. But the lack of scientific research on them means existential risk from climate change remains poorly understood.

Several of the greatest catastrophes in human history have been caused by pandemics. The Black Death of 1347 killed 25–50 percent of people in Europe—about a tenth of the world's population.[11] The introduction of diseases from Europe (beginning in 1492) may have killed as much as 90 percent of the population in the Americas—again about a tenth of the world's population.[12] The 1918 flu killed roughly 3 percent of the world's population.[13]

So the current worldwide pandemic is not at all unprecedented. It is the worst pandemic in a century, but far from the worst in a millennium. Indeed, it is the idea that such catastrophes were left forever behind us that would have been unprecedented. Covid-19 shows us that this is false, that humanity is still vulnerable to global catastrophes. While we have made substantial improvements in medicine and public health (which have greatly reduced the burden of endemic disease), it is unclear whether we are any safer from pandemics. This is because there are also ways that human activity has made pandemics more dangerous, such as intensive farming, urbanization and rapid international travel. So even when pandemics are natural in origin, the argument for bounding natural extinction risk does not apply—that argument assumes the risk has been stable or declining over human history, which may not be true here. Though Covid-19 itself does not pose an existential risk to humanity, other pandemics might.[14]

And this situation looks considerably worse when we consider the possibility of engineered pandemics. Humanity has a long and dark history with using disease as a weapon, dating back at least 3,000 years.[15] Indeed, there are credible claims that the Black Death was introduced into Europe by catapulting plague-ridden bodies into the besieged city of Caffa on the Crimean Peninsula.[16] The 20th century saw many countries adopt major biological weapons programmes, and while these were officially outlawed by the Biological Weapons Convention of 1972, it would be a serious mistake to think that the convention has stopped all bioweapons programmes.[17] Though it is an important symbol and a useful forum, it is very under-resourced: with just four employees and a budget smaller than that of a typical McDonald's.

Biotechnology is advancing at an extremely rapid rate. And while these advances bear great promise for medical and industrial progress, they also aid progress in biological weaponry. This makes the weapons of a major state more powerful and opens up the possibility of extremely damaging weapons being deployed by small nations or subnational groups. If biotechnology continues to advance, this may create a very unstable strategic situation.

And there are other important technological risks on the horizon, such as those posed by advanced artificial intelligence and nanotechnology.[18] The sheer variety of these risks suggests that a piecemeal, siloed, approach—in which we hope that each risk will be dealt with separately by the relevant community—becomes increasingly hard, and a more unified approach is needed.

The anthropogenic risks are inherently more speculative than the natural risks, since it is impossible to acquire evidence of them having happened before. But this does not make them smaller. We saw that natural risk almost certainly totals less than 1 in 300 per century. How confident would we be that humanity could expect to survive 300 centuries like the 20th century? Or like the 21st? Using the fossil record, we can be more than 99.7 percent confident we will survive the natural risks of the next 100 years. How confident can we be that we survive the human-made risks? While we cannot be sure, reflections such as this make it seem likely that anthropogenic risks are now the greater threat to our future, posing an unsustainable level of risk (box S1.3.1).

Analysis

The world is only just beginning to understand the scale and severity of existential risk. The substantial work on the risks of nuclear war and climate change still pales in comparison with the importance of the topics. And little of this work has been directed to the parts of these problems most relevant to existential

Box S1.3.1 Existential risk as sustainability

Protecting humanity's long-term potential is a key form of sustainability. The current period of heightened anthropogenic risk is unsustainable—we can get lucky for a while, but eventually the odds are going to catch up with us. In many other cases people can do well by taking calculated risks, but here our entire bankroll is on the line, so if we eventually lose—even once—there is no coming back.

We could thus think of our accumulated existential risk over humanity's future as a kind of risk budget—a budget that has to last for our entire lifespan, the ultimate nonrenewable resource. Responsible stewardship of humanity's potential would involve lowering this risk as quickly as possible and setting in place the safeguards to keep it low in order to allow humanity to flourish for as long as possible.

risk (such as better understanding nuclear winter or extreme climate feedbacks).

It is helpful to look at why existential risk is so neglected.

First, protection from existential risk is an intergenerational global public good. Standard economic theory thus predicts a market failure in which individual nations cannot capture more than a small fraction of the benefits and are tempted to free-ride on each other, undersupplying this protection.

Second, many of the risks are inherently international—beyond any individual nation's ability to solve, were one even prepared to do so. International cooperation and coordination are thus required but move much slower than technology. If we remain in a paradigm in which a new agreement is required for each new risk and can be achieved only decades after the risk rises to prominence, we might forever be playing catchup.

Third, minimizing existential risk just feels like too big a task for most nations—something that is outside the scope of their usual responsibilities or "above the pay grade" of their leaders. Yet nations have not officially passed this responsibility up to the international level, entrusting an international institution with key tasks relating to monitoring, assessing or minimizing existential risks. Responsibility for protecting humanity's long-term potential thus falls through the cracks between the national and international spheres.

Fourth, the whole idea of existential risks to humanity is very recent. We have been exposed to anthropogenic existential risks for only 75 years, most of which was spent in the grip of a Cold War. Our ethics and our institutions have not had time to catch up.

As we begin to wake up to the present situation, we will face great challenges. But there will also be new opportunities. Responses that first seemed impossible may become possible—and in time even inevitable. As Ulrich Beck put it, "One can make two diametrically opposed kinds of assertion: global risks inspire paralysing terror, or: global risks create new room for action."[19]

We have seen that the rise in anthropogenic risk means that most of the existential risk we face likely arises from our own actions. While this is a disturbing trend, there is a flip side that should give us hope: Humanity's future is largely within humanity's control. If a 10 kilometre across asteroid were on a trajectory to hit the Earth in 10 years, there might truly be nothing we could do to stop it. But the risks from nuclear war, climate change and engineered pandemics arise from activities that humans perform—and thus that humans can stop.

There are serious challenges to doing so—challenges of international coordination, verification and policing—as well as the overarching challenge of creating the political will for action. But these are not insurmountable.[20] If we fail, it will not be because there was no way through but because we were distracted by other issues or were not willing to do the necessary work. If we set our minds to it, taking the risks with due seriousness and adopting the protection of humanity's long-term potential as one of the overarching missions of our time, then our generation could very well be the one that sets humanity on a path towards a long, secure future.

NOTES

1 Ord 2020.

2 The idea of existential risk was introduced by Bostrom (2002). Earlier work on the ethics of human extinction includes Leslie (1996), Parfit (1984), Sagan (1983) and Schell (1982).

3 Longrich, Scriberas and Wills 2016.

4 Stokes and others 2017.

5 See also Snyder-Beattie, Ord and Bonsall (2019).

6 Sagan 1983.

7 Robock, Oman and Stenchikov 2007.

8 For example, Richard Turco (Browne 1990): "My personal opinion is that the human race wouldn't become extinct, but civilization as we know it certainly would." And Alan Robock (Conn, Toon and Robock 2016): "Carl [Sagan] used to talk about extinction of the human species, but I think that was an exaggeration. ... But you wouldn't have any modern medicine. ... You wouldn't have any civilization."

9 Hertsgaard 2000; Reagan 1985.

10 See Ord 2020.

11 See Ord 2020.

12 See Ord 2020.

13 Taubenberger and Morens (2006) estimate 50 million deaths, which would be 2.8 percent of the 1918 world population of 1.8 billion.

14 Snyder-Beattie, Ord and Bonsall 2019.

15 Trevisanato 2007.

16 Kelly 2006.

17 Countries that are confirmed to have had bioweapons programmes include Canada (1940–1958), Egypt (1960s–?), France (1915–1966?), Germany (1915–1918), Iraq (1974–1991), Israel (1948–?), Italy (1934–1940), Japan (1934–1945), Poland (?), Rhodesia (1977), South Africa (1981–1993), Soviet Union (1928–1991), Syrian Arab Republic (1970s?–?), United Kingdom (1940–1957) and United States (1941–1971). See Carus (2017).

18 For more on existential risk from artificial intelligence, see Bostrom (2014) and Russell (2019). For existential risk from nanotechnology, see Drexler 2013.

19 Beck 2009, p. 57.

20 For a list of concrete policy and research proposals that would make a difference, see Ord (2020).