

# 장기주의에 반대한다

원본 - <https://aeon.co/essays/why-longtermism-is-the-worlds-most-dangerous-secular-credo>

저자 - Émile P Torres

편집 - Sam Dresser

역자 - 김유민(daseoh@gmail.com)

장기주의는 변두리에서 시작한 인류의 미래에 대한 철학 연구였지만, 이제는 많은 경제적 지원을 받고 있고, 그 위험성 또한 높아지고 있다.

---

인간이 종말에 다다를지도 모른다는 인식은 높아지고 있는 것 같다. 재앙에 대한 끔찍한 예측은 언론을 어지럽힌다. 지옥 같은 산불이나 위력적인 홍수, 코로나 환자로 넘치는 병원에 대한 영상이 우리 타임라인을 점령한다. 멸종에 저항하는 운동가들은 세계를 구하기 위해 필사적으로 도시를 봉쇄하고 있다. 인류의 미래에 대해 묻는 한 조사에서 절반이 넘는 응답자는 우리 삶의 방식이 백 년 안에 끝날 가능성이 50 퍼센트가 넘는다고 답했다.

종말론이나 종말이 가까웠다는 믿음은 물론 새로운 것이 아니다. 사람들은 수천 년 동안 종말이 다가왔다고 경고했다. 사실, 많은 신학 연구자들은 예수가 그의 삶 동안 종말을 기대했다고 믿는다. 그러나 오늘의 상황은 과거와 근본적으로 다르다. 현재 논의되는 종말론 시나리오는 종교적 예언자의 폭로나 마르크시즘 같은 인간 역사에 대한 비종교적인 거대 담론이 아니라 기후학이나 생태학, 유행병학의 유명한 전문가들이 내놓은 확고한 과학적 결론으로 뒷받침된다.

예를 들어서, 우리는 기후 변화가 문명에 끔찍한 위협이 되는 자리에 올 것을 안다. 우리는 생물다양성 상실과 여섯 번째 대규모 멸종이 지구 생태계에 갑작스럽고 되돌릴 수 없는 재앙적 변화를 촉진할 수 있다는 것을 안다. 핵전쟁은 태양을 몇 년에서 몇십 년간 가릴 것이다. 이는 전 세계 농업을 파괴할 것이다. 그리고 SARS-CoV-2 가 우한 연구소에서 나왔든 자연의 주방에서 만들어 졌든(아마도 후자가 더 그럴듯하지만), 합성 생명학은 곧 다원적 진화가 만들 수 있는 그 어떤 것보다 치명적이고 전염력이 강한 병원체를 만드는 역할을 하게 될 것이다. 철학자들과 과학자들은 인공 지능, 분자 나노공학 그리고 성층권 지구공학에 관련한 새로운 위협에 대한 경종을 울리기 시작했지만, 이는 그다지 강하지 않아 보인다.

이런 고려는 많은 학자들이 2016 년에 스티븐 호킹이 <더 가디언>에서 "우리는 인류 발전의 가장 위험한 순간에 있다"라고 한 말을 인정하게 했다. 예를 들어서, Martin Rees 경은 인류가 2100 년에 도달할 확률은 반반이라고 예측했다. Noam Chomsky 는 절멸의 위험은 현재 '호모 사피엔스의 역사에서' 전례 없는 것이라고 주장했다. 그리고 Max Tegmark 는 "아마도 우리가 사는 동안 ... 우리가 자멸하거나 정신 차리게 될 것이다"라고 말했다. 이런 우울한 선언에 맞춰서, 2020 년에는 *Bulletin of the Atomic Scientists* 의 상징인 세계 종말 시계가 자정(혹은 종말)까지 100 초 남은 시점을 가리키게 되었다. 이는 이 시계가 만들어진 1947 년 이후 자정과 가장 가까운 것이다. 그리고 2020 년에 11,000 명이 넘는 전 세계 과학자들은 "분명하고 명확하게 지구가 기후 위기에 처했다"라고 말하면서 "생물권을 보존하기 위한 노력의 규모를 엄청나게 늘리지" 않으면 "기후 위기로 인해 지금까지 겪어보지 못한 고통을" [겪을 것이]라는 내용의 기고문에 서명했다. 2019 년에 젊은 기후 활동가인 Xiye Bastida 가 *Teen Vogue* 에서 말한 이런 존재에 관한 분위기를 요약한 것에 따르면, 그 목적은 "우리가 마지막 세대가 되지 않도록" 하는 것이다. 왜냐하면 마지막 세대가 될 가능성이 상당히 때문이다.

오늘날 인류가 마주한 전례 없는 위험을 고려하면, 철학자들이, 인류 멸종에 관한 윤리적 함의나 문명의 영구적 붕괴 같은, 이와 연관된 시나리오에 상당한 양의 잉크를 사용했다고 추정할 수 있다. 우리가 사라지는 것이 윤리적으로 얼마나 (좋거나) 나쁠까? 그리고 어떤 이유로 사라지는 것일까? 미래 세대가 존재하지 못하게 하는 것이 잘못된 것일까? 과거의 희생, 투쟁 그리고 노력의 가치가, 지구 혹은 더 넓게는 우주가 거주할 수 있는 상태로 유지되는 한, 인류가 계속 존재하는 것에 달려있을까?

그러나 철학자들이 이 문제에 대해 노력했다는 예상은 사실이 아니다. 지금까지 철학자들은 우리의 멸종이라는 주제에 대해 거의 신경 쓰지 않았고, 심지어 지금도 철학적 논의의 변두리에 위치하고 있다. 전체적으로 철학자들은 다른 주제에 열중하고 있다. 그러나 여기에 한 가지 예외가 있다. 지난 이십 년간, 대부분 옥스퍼드에 기반한 작은 학자 집단은 장기주의(longtermism)라고 불리는 새로운 도덕적 세계관의 세부 사항을 해결하는 일에 골몰해 왔다. 장기주의는 우리 행동의 결과가 아주 오랜 기간에 걸쳐서, 그러니까 지금으로부터 몇천, 몇백만, 몇억, 심지어 몇조 년 뒤에, 미칠 영향을 강조한다. 장기주의는 Nick Bostrom 과 Nick Beckstead 의 작업에서 시작한다. 2005 년에 Bostrom 은 이름도 거창한 인류 미래 연구소(Future of Humanity Institute)를 설립했고 Beckstead 는 FHI 의 연구원이다. 그리고 주로 Toby Ord 에 의해 공개적으로 옹호되었다. Ord 는 Open Philanthropy 의 프로그램 책임자이고 <<사피엔스의 멸망>>의 저자다. 장기주의를 주요 주제로 삼는 연구소에는 Global Priorities Institute(GPI)와 Forethought Foundation 이 있다. GPI 는 Hilary Greaves 가 운영하고 FHI 와 관계가 있다. 그리고 FF 는 FHI 와 GPI 에 자리를 두고 있는 William MacAskill 이 운영한다. 직함, 이름, 기관 그리고 약어가 복잡하게 얽혀있는 것에 더해서, 장기주의는 효과적 이타주의 운동(Effective

Altruism movement)의 주요 주제(?, cause areas) 중 하나다. 2011년에 Ord가 도입한 EA는 지금은 460억 달러라는 큰 규모의 자금 확약을 확보하고 있다.

장기주의의 영향력은 과장하기 어렵다. 1845년에 Karl Marx는 단지 세계를 해석하는 것이 아니라 세계를 변화시키는 것에 철학의 방점이 있다고 선언했다. 그리고 이것이 장기주의자들이 비범한 성공을 거두며 해왔던 일이다. Bostrom의 작업을 지지해 온 Elon Musk가 더 거창한 이름의 자매기관인 생명의 미래 연구소(Future of Life Institute)를 통해 15억 달러를 FHI에 기부한 것을 생각해보자. Jann Tallinn은, 최근 내가 다른 글에서 말했듯이, 장기주의 이데올로기를 거수하며 기후 위기가 인류에 존재적 위기가 될 수 없다고 믿는 백만장자 기술 사업가인데, FHI의 공동 설립자이기도 하다.

한편, Peter Thiel은, EA 콘퍼런스에서 기초연설을 했고 억만장자 자유주의자들과 도널드 트럼프를 지지하는 인물인데, 많은 돈을 Machine Intelligence Research Institute에 기부했다. 이 단체의 목표는 초지능 기계로부터 인류를 구원하는 것이다. 이 목표는 장기주의와 깊은 관련이 있다. GPI나 Forethought Foundation와 같은 다른 단체들도 젊은 사람들을 끌어들이기 위해 에세이 경진 대회와 장학금을 지원한다. 또 워싱턴 DC에 자리하고 있는 Center for Security and Emerging Technologies가 장기주의자를 국제 정책을 다루는 미국 정부 고위직에 놓는 것을 목표로 삼고 있다는 것은 공공연한 사실이다. CSET의 설립자인 Jason Matheny는 과거에 FHI 연구원이었고 지금은 미 대통령 Joe Biden 아래서 기술과 국제 안보를 담당하는 부보좌관(?, deputy assistant)으로 일하고 있다. Ord는, 철학자치고는 놀랍게도, World Health Organization, World Bank, World Economic Forum, 미국의 National Intelligence Council, 영국의 Prime Minister's Office, Cabinet Office, Government Office for Science에서 조언하는 역할을 맡고 있다. 또 구체적으로 장기주의(long-termism)을 언급한 UN의 사무총장 명의의 보고서를 만드는 일에 기여했다.

요점은 상위 대학과 실리콘 벨리 밖 사람들은 거의 모르는 장기주의가 가장 영향력 있는 이념 중의 하나가 될 수 있다는 것이다. 나는 변화가 필요하다고 믿는다. 왜냐하면 나는 과거에 장기주의자였고, 4년 전에 이 개념을 옹호하는 책 전체를 집필한 경험이 있는데, 이 세계관은 오늘날에 꽤 가능성 있으면서 정말 위험한 비종교적 신념 체계로 생각하기 때문이다. 일단 짐승의 특성을 이해하려면 먼저 그 해부학적 특성과 생리적 기능을 이해해야 한다.

---

우선, Bostrom과 Beckstead가 제안한 장기주의는 장기적 관점을 중요시하거나 미래 세대의 복지를 중요하게 여기는 것과 다르다는 것에 주목해야 한다. 장기주의는 그 이상이다. 장기주의의 주제는 단순하다. (비록 결함이 있지만, 내 생각에는) 개인과 인류 전체 사이의 유추다. 이 생각을 이해하기 위해서, 동료들이 자기 세대에서 뛰어나다고 평가하는 케임브리지 교수 Frank Ramsey의 사례를 생각해보자. Belletrist Lytton

Strachey 는 Ramsey 에게 뉴턴의 어떤 것이 있다고 말한 적이 있다. G E Moore 는 Ramsey 의 "매우 뛰어난 탁월함"에 대해 쓰기도 했다. 그리고 John Maynard Keynes 는 Ramsey 의 글을 "수리 경제학 사상 가장 놀라운 기여 중 하나"라고 평가했다.

그러나 Ramsey 의 삶은 행복하지 않았다. 1930 년 1 월 19 일에 그는 런던 병원에서 수술 후에 죽었는데, 케임브리지를 가로지르는 캠 강에서 수영 중에 간염이 감염된 것이 가장 의심되는 사인이다. 이때 겨우 26 살이었다.

이 결과가 비극인 이유에는 명확한 두 가지 이유가 있다고 주장할 수 있다. 첫 번째 이유는 분명하다. Ramsey 의 삶이 빨리 끝났고, 그가 살아있었다면 경험할 수 있었을, 즐거움과 행복, 사랑과 우정 같은 인생을 가치 있게 만들, 모든 것을 빼앗겼기 때문이다. 이 관점에서 Ramsey 의 이른 죽음은 개인적 비극이다. 그러나 두 번째로, 그의 죽음은 인류 지식에 엄청난 기여를 할 운명이었던 지식계의 슈퍼스타를 세계에서 빼앗았다. Partha Dasgupta 경은 "Ramsey 가 연 길은 많다"라고 말했다. 그렇지만 얼마나 많은 길이 열릴 수 있었을까? Strachey 는 The loss to your generation is agonising to think of what a light has gone out"라고 한탄했다. 이는 만약 Ramsey 가 요절하지 않았다면 서양의 지적 역사가 얼마나 다르게 변했을지 궁금하게 만든다. 이 관점에서는, 비록 Ramsey 의 죽음이라는 개인적 비극이 정말로 참담하지만, 세계를 더 낫게 만들 그의 광대한 잠재력이 더 심각한 두 번째 비극을 낳는다. 달리 말하면, 이 죽음의 나쁨은, 대부분, 아마도 압도적으로, 그가 경험한 개인적이고 직접적인 피해에서 나오는 것이 아니라, 실현되지 못한 잠재력에서 나온다. 혹은 그렇게 주장할 수 있다.

장기주의자는 마치 인류가 일생동안 낭비, 성취, 파괴, 실현할 수 있는 잠재력이 있는 한 개인인 것처럼, 앞서 주장과 결론을 인류 자체에 대입한다. 한편으로, 인구를 0 으로 줄일 수 있는 재앙은 당시 살아있는 사람들에게 가하는 모든 고통 때문에 비극일 수 있다. 핵전쟁 이후 수년에서 수십 년이 지난 시기에 정오이지만 칠흑 같은 하늘 아래 영하의 기온을 느끼며 굶어 죽는 공포를 상상해 보자. 이것은 직접 영향받을 사람들에게 개인적인 고통인, 그러니까 첫 번째 부류의, 재앙이다. 그러나 장기주의자의 주장에 따르면, 두 번째 비극은 첫 번째보다 천문학적으로 나쁘다. 왜냐하면 인류의 절멸은 앞으로, 말하자면 10 의 100 승 년 정도로 길게 (열평형 상태가 되어서 생명이 불가능할 시점까지) 번영할 수 있는 미래를 앗아갈 것이 분명하기 때문이다. 그렇다면 인류의 절멸은, Ord 의 거의 종교적인 표현에 따르면, 인류의 "광대하고 영광스러운" 장기적인 잠재력을 파괴하는 것이다. 우주의 규모와 열역학적 평형에 이를 때까지 남은 시간을 고려하면, "잠재력"은 거대해서 첫 번째 비극은 거의 보이지 않을 것이다.

이는 곧바로 개인과 인류 사이에 또 다른 유사점을 제시한다. 죽음이 누군가의 잠재력이 실현되지 못하는 유일한 원인은 아니다. Ramsey 가 일찍 죽지 않지만 연구하고 글 쓰고 학문적인 글을 출판하는 대신에 동네 술집에서 당구 치고 술 마시는 경우를 생각해보자. 다른 실패 방법이지만 같은 결과다. 장기주의자들은 이를 인류에 적용해서 우리가 죽지 않아도 실패할 수 있는 방법이 있다고 주장할 것이다. 아래에서 다시 이 지점으로 돌아올 것이다.

## 이 관점에서 기후 위기는 아흔 노인 발가락에 있는 돌 때 생긴 흉터에 불과하다.

지금까지 논의를 정리하면, 인류는 각각의 개인의 잠재력을 초과하는 잠재력을 가지고 있고, 이 잠재력을 실현하는 데에 실패하는 것은 극도로 나쁜 것이다. 앞으로 살펴보겠지만, 나쁜 수준이 아니라 문자 그대로 우주적 수준의 도덕적 재앙이다. 이것이 장기주의의 핵심 교리다. 윤리적으로 보면, 지구 기반의 지적 생명인 우리의 잠재력을 실현시키는 것보다 더 중요한 것은 없다. 이 문제는 중요한데, 심지어, 장기주의자들은 "존재적 위험"(existential risk)과 "존재적 재앙"이라는 무시무시한 단어를 만들 정도다. 존재적 위험은 우리 잠재력을 파괴할 어떤 가능성을 의미하고, 존재적 재앙은 실제로 이 잠재력을 파괴할 어떤 사건을 가리킨다.

내가 장기주의를 위험하다고 생각하는 이유는 무엇일까? 간단히 대답하면, 인류의 잠재력 실현을 최고의 가치로 여기는 것은 오늘날 혹은 가까운 미래에 살 인간이 극도의 위험을 겪거나 죽을 확률을 크게 높인다. 내가 다른 곳에서 지적했지만, 장기주의 이데올로기 추종자들이 기후 변화에 무심한 태도를 갖는 경향이 있다는 것을 생각해보자. 왜 그럴까? 그 이유는, 기후 변화가 섬 나라들을 사라지게 하고 대규모 이민과 백만이 넘는 사람들의 죽음을 만든다 해도, 기후 변화는 다가오는 수조 년의 장기적인 잠재력과 타협할 수 있는 대상이 아니기 때문이다. 기후 재앙이 앞으로 이 천 년 동안 나올 인구의 75 퍼센트를 감소시킨다 해도, 이 문제에 대해 우주적 관점을 가지고 있는 사람들에게는 작은 사건에 불과하다. 마치 아흔 노인 발가락에 있는 돌 때 생긴 흉터에 불과하다.

Bostrom 의 주장에 따르면 인류 전체의 관점에서 보았을 때 전 세계 문명의 파괴를 유발할 존재적이지 않은 재앙은 복원 가능성이 있다. 그리고 인류에 대한 대학살이 될 수 있지만, 인류가 그 잠재력을 실현하려고 다시 회복하는 한, 인류의 작은 실수 정도로 기록될 것이라고 덧붙인다. 한편, Bostrom 은 거대한 관점에서 보았을 때, 역사상 최악의 자연재해와 지독한 잔혹성은 거의 알아차리기 힘든 사소한 것이 된다고 말한다. Bostrom 은 두 번의 세계 전쟁과 에이즈, 체르노빌 방사능 유출 사건을 언급하면서 "이런 비극은 사람들에게 즉시 영향을

주지만, 큰 그림으로 보면 ... 심지어 이런 재앙 중 가장 최악의 사건도 삶의 큰 바다에 이는 물결에 지나지 않는다"라고 선언한다.

에이즈와 홀로코스트의 나뭇을 평가하는 것처럼 세상을 보는 방식은 동일한 (존재적이지 않은) 정도와 강도를 지니는 다가올 재앙도 단지 물결로 구분될 수밖에 없음을 암시한다(?). This way of seeing the world, of assessing the badness of AIDS and the Holocaust, implies that future disasters of the same (non-existential) scope and intensity should also be categorised as mere ripples). 만약 이런 재앙이 직접적인 존재적 문제가 아니라면 크게 걱정할 필요가 없다. 다만 개인에게 해당하는 비극일 뿐이다. 2003 년에 Bostrom 은 "네 번째 우선 순위까지는 ... 존재적 위험을 감소하는 데 사용되어야 한다"라고 썼다. 몇 년 후에 그는 한정된 자원을 세계 빈곤 문제를 경감하거나 동물의 고통을 줄이는 데 낭비해서는 안 된다는 주장을 반복했다. 이 둘 모두 우리의 장기적 잠재력을 위협하지 않는데, 장기적 잠재력이 정말로 중요한 것이기 때문이다.

Ord 는 이 관점에 따라서 인류의 가장 중요한 과제는 ... 안전한 공간에 들어가는 것이다"라고 주장하면서 이 공간을 존재적 위험이 낮으면서 그 낮은 상태를 유지하는 곳"이라고 정의한다. 그리고 이 상태를 존재적 안전"이라고 부른다. 무엇보다도 중요한 것은, 즉각적인 위험을 피하고 인류가 실패하는 일이 불가능하도록, 인류를 장기적 미래에 걸친 위험에서 보호하는, 믿음만한 안전 요원을 고안해서, 우리의 잠재력을 보존하고, 이를 위해 할 수 있는 모든 것을 다 해야 한다는 것이다. 비록 Ord 는 기후 변화에 대해 동의하지만, 의심스러운 방법론에 기대서, 한편으로는 기후 변화가 존재적 비극을 유발할 확률이 1000 분의 1 에 불과하다고 주장한다. 이 확률은 Ord 에 따르면 이번 세기에 초지능 기계가 인류를 파괴할 확률보다 전체적으로 두 자릿수 이상 낮은 수준이다.

여기서 정말로 주목해야 할 것은 (Bostrom 의 언어로, 넓은 관점에서 보았을 때, 인류에게 기후 재앙은 단지 작은 실수일 뿐이라는 것을 기억하자) 기후 재앙이 전 세계 사람들에게 미치는 영향 아니라, Ord 가 <<시피엔스의 멸망>>에서 말했듯이, 기후 재앙이 문명에 회복 불가능한 위험이 되거나 인류를 완전히 멸종시킬 가능성이 희박하다는 점이다. 다시 말하자면, (특히 남반구) 사람들에게 미치는 피해는 절대적인 측면에서 상당한 의미가 있지만, 우주에서 펼쳐질 장기적인 잠재력의 광대함과 영광스러움을 비교해보면, 이런 피해는 거의 보이지 않는다.

---

장기주의의 함의는 훨씬 더 걱정스럽다. 만약 존재적 재앙을 피하는 (즉 우리 잠재력을 실현하는) 일을 위해 인류의 최우선 과제 네 가지를 할당한다면, 무엇을 선정해야 할까? 더 거대한 선이라고 여기는 개념이 (전쟁 같은) 잔인성을 정당화하는 방법에 대해 Nagel 이 말한 것을 생각해보자. 그는 만약 목적이 방법을 정당화하고

목적이 (국가 안전 같이) 충분히 중요하다고 여겨지면, "이는 불탄 아이들에 대한 책임이 있는 사람들의 양심의 가책을 덜어 줄 수 있다"라고 설명한다. 이제, 더 큰 선이 국가 안전이 아니라 앞으로 수 조년 동안 지구 기반 지적 생명의 우주적 잠재력에 관한 것이라면, 무엇이 정당화 될 수 있는지 생각해보자. 이차 세계 대전 동안 4 억 명이 넘는 시민이 희생당했지만, 이 숫자를 (Bostrom의 추정대로) 우리가 존재적 재앙을 피할 수 있어서 나올 수 있는 10의 54승 이상인 인구에 비교해보자. 이 잠재력을 보호하거나 보존하기 위해 우리가 하지 말아야 할 것은 무엇인가? 아직 태어나지 않은 이런 사람들의 존재를 보장하기 위해서? 우주적으로 중요한 도덕적 목적으로 정당화할 수 없는 것은 무엇일까?

Bostrom은 예방적 치안 능력을 강화하기 위해 (즉, 문명을 파괴할 수 있는 대량 학살 테러리스트를 예방하기 위해) 전 세계를 대상으로 모든 사람을 실시간으로 관찰하는 침습적 감시 시스템을 만드는 것을 진지하게 고려해야 한다고 주장했다. 또한 존재적 재앙을 피하기 위해 국가가 선제적인 폭력 또는 전쟁을 활용해야만 한다고 말했다. 그리고 수십억의 사람을 구하는 것과 존재적 위험을 아주 작은 수준으로 감소시키는 것은 도덕적으로 동등하다고 주장했다. 그의 말에 따르면, 비록 미래에 10의 54승의 사람이 존재할 1퍼센트의 확률이라도 있다면, 존재적 위험을 단지 **1 퍼센트포인트의 십억분의 일의 십억분의 일**로 줄이는 것은 십억 명이 사는 것보다 천억 배 가치 있다. (일부 장기주의자들도 인정하는) 이런 광신주의 때문에 만약 정치 지도자들이 실제 세계에서 Bostrom의 관점을 진지하게 여기면 무슨 일이 일어날지 걱정하는 비평가들이 많아지고 있다. 다음은 (당혹스럽게도) 다른 방식으로 장기주의를 호의를 갖고 말하는 수리 통계학자 Olle Häggström의 말이다.

나는 정치인들과 의사 결정자들이 (위에서 계산한 것을) 문자 그대로 받아들일 가치가 있는 정책의 가이드로 여기는 관점에 대해 극도의 불안감을 느낀다. 이는 단순히 "오물렛을 만들려면, 기꺼이 계란 몇 개를 깨야 한다"라는 오래된 속담을 너무 연상시킨다. 이 속담은 일반적으로 미래의 유토피아를 만드는 과정에 필요하다면, 약간의 집단학살 정도는 선한 것이 될 수 있다고 설명할 때 쓰인다. CIA의 수장이 미국 대통령에게 종말을 가지고 올 무기를 만들어서 인류를 쓸어버릴 미치광이가 독일 어딘가에 있다는 확실한 증거가 있고 이 미치광이가 성공할 확률이 십억분의 일이라고 설명하는 상황을 상상해 보자. 이들은 이 미치광이가 어디 있는지 또는 누구인지에 대한 더 이상의 정보가 없다. 만약 대통령이 Bostrom의 주장을 진정으로 받아들이고 확률 계산을 할 줄 안다면, 그는 독일 국경 안에 모든 사람들을 죽이기 위해 총력적인 핵 공격을 감행하는 것이 가치 있는 일이라고 결정할 것이다.

그러니까, 여기에 장기주의가 근원적으로 위험해질 수 있다고 생각하는 몇 가지 이유가 있다. 그러나 장기주의 세계관에는, 내가 아는 바로는, 지금까지 아무도 지적하지 않은 추가적인 심각한 문제도 있다. 예를 들어, 장기주의에 대한 근본적인 믿음이 인류가 당장 생존에 대해 예측할 수 없는 많은 위험에 처하게 된 중요 이유로 여길 좋은 사례가 있다. 다른 말로, 장기주의는 존재적 안전을 달성하는 것과 양립하지 않는다. 이는 인류가 미래에 멸종이나 붕괴할 가능성을 진정으로 낮추는 유일한 길은 장기주의 이데올로기를 완전히 포기하는 것임을 의미한다.

## Bostrom 과 Ord 에 따르면 포스트 휴먼이 되는 것에 실패하는 것은 인류가 광대하고 영광스러운 잠재력을 현실화하지 못하게 된다는 것이다.

이 문장을 이해하기 위해서, 우선 지금까지 정의 없이 사용한 "장기적 잠재력"(longterm potential)이 장기주의자들에게 무엇을 뜻하는지 알아보자. 이 개념은 포스트휴머니즘, 우주 확장주의 그리고 전체 공리주의의와 상당히 관련된 윤리적 관점, 이렇게 세 부분으로 분석할 수 있다.

첫 번째는 발전된 기술로 우리 몸과 뇌를 개량해서, 근본적으로 강화된 (혼란스럽게도 장기주의자들은 인류로 구분하는) 포스트휴먼에 속하는, 우월한 종족을 만들어야 한다는 생각이다. 비록 Bostrom 이 가장 저명한 트랜스휴머니스트이지만, 장기주의자들은 트랜스휴머니즘이라는 용어를 사용하기 꺼린다. 아마도 부정적인 생각이 연상되기 때문일 것이다. 예를 들어, Susan Levin 은 현대 트랜스휴머니즘이 백인 우생학 운동에서 나왔다고 지적한다. 그리고 Bostrom 과 함께 *Human Enhancement* (2009)를 집필한 Julian Savulescu 같은, 트랜스휴머니스트들은 (그는 "ultimate harm"이라고 하는) 존재적 재앙을 막기 위해, 옥시토신 같은, 문자 그대로 도덕성 고취 약물을 사용할 것을 주장한다. Savulescu 는 동료와 함께 "인간 도덕성을 향상시키는 것은 시급한 문제이고 ... 우리는 이를 위해 어떤 방법이라도 찾아야 한다"라고 했다. 이러한 주장은 논쟁적일 뿐만 아니라 많은 사람들에게 매우 불안한 것이다. 따라서 장기주의자들은 장기주의를 옹호하지만 이런 생각에서 거리를 두기 시작했다.

트랜스휴머니즘은 지금 인간 방식보다 더 나은 존재의 다양한 포스트휴먼 방식이 있을 것이라고 주장한다. 예를 들어, 우리는 유전적 방법을 통해 우리 감정에 대한 완전한 조절이 가능한 존재가 될 수 있다. 또 신경에 삽입한 장치로 인터넷에 접속하거나, 심지어 디지털 영생을 위해 마음을 기계에 업로드할 수도 있다. Ord 가 <<사피언스의 종말>>에서 주장한 것처럼, 붉은 여우나 비둘기가 갖고 있는 지자기 감각이나 박쥐나 돌고래처럼 방향을 방향 위치 감각을 통해, 세계를 지각하는 것이 얼마나 멋진 일인지 생각해보자. Ord 는 "이런

미지의 경험들은 우리보다 정교하지 않은 의식들 속에 있다. 그렇다면 엄청난 가치를 지닐 수 있는 경험 중에 어떤 것이 더 위대한 의식에 접근할 수 있을까"라고 말한다. 이 가능성에 대한 Bostrom 의 가장 기상천외한 탐구는 그의 미래 암시적인 글인 <유토피아에서 온 편지>에서 나온다. 이 글 속에서는 가상의 포스트 휴먼이 "우리는 즐거움을 음료에 섞습니다"라고 말할 정도이다. 이 정도로 즐거움이 넘쳐나는 초지능 포스트휴먼들로 가득 찬 기술 유토피아 세계가 묘사되는 것이다.

Bostrom 과 Ord 에 따르면, 장기주의와 트랜스휴머니즘의 관계는 포스트 휴먼이 되는 것에 실패하는 것은 인류가 광대하고 영광스러운 잠재력을 현실화하지 못하게 된다는 것이다. 이 상황은 곧 존재적 재앙이다. 2012 년에 Bostrom 은 "인간의 생물학적 본성을 변화시킬 가능성을 영구적으로 상실하는 것은 존재적 재앙이다"라고 썼다. 유사하게 Ord 는 "지금의 인류를 영원히 보존하는 것은 우리 유산을 낭비하고 우리 잠재력의 대부분을 포기하는 것과 같을 수 있다"라고 주장했다.

---

우리 잠재력의 두 번째 부분인 우주 확장주의는 이론적으로 접근 가능한 시공간인 미래의 빛 원뿔 영역을 가능한 많이 식민지화 해야 한다는 생각을 말한다. 장기주의자들에 따르면 우리 미래의 빛 원뿔 영역에는 엄청난 양의 이용 가능한 자원이 있다. 이 자원을 장기주의자들은 반 엔트로피 (혹은 반대 엔트로피)의 우주적 자산이라고 부른다. Ord 는 은하만 해도 150,000 광년에 걸쳐서, 대부분 자체 행성을 가지고 있는 1,000 억 개의 별을 포함하고 있다고 말한다. 그는 계속해서, 인류의 장기적 잠재력을 달성하기 위해 [우리는] 결국 이웃 행성으로 가서 새로운 풍부한 사회를 만들기 위한 발판을 충분히 다져야 하고 또 여기서 더 멀리 모험할 수 있을 것이라고 말한다. 한 번에 6 광년씩만 나아간다면, 포스트 휴먼 후손은 우리은하의 거의 모든 별에 도달할 수 있을 것이다. 왜냐하면 은하계 전체를 생명으로 채우기 위해서 우리 항성계를 포함해서 각 항성계에 가장 가까운 별 몇 개에만 정착하면 되기 때문이다. 이 과정은 기하급수적으로 진행될 수 있고, 우리 후손이 별에서 별로 이동하는 때 순간마다 점점 더 번영하는 사회를 만들 것이다.

그러나 정확히 왜 이것이 필요할까? 새로운 포스트 휴먼 문명으로 우주가 채워지는 것에서 중요한 것이 무엇일까? 이 질문은 잠재력의 세 번째 부분으로 이끈다. 바로, 내가 짧게 공리주의라고 부르는 전체 공리주의다. 비록 일부 장기주의자들은 자신이 공리주의자가 아니라고 주장하지만, 이는 대부분의 장기주의, 더 일반적이게는, 장기주의가 묻어나는 효과적 이타주의(Effective Altruism) 운동이 공리주의의 재포장에 불과하다는 비판을 왜곡하기 위한 연기와 거울 전략이라는 것을 바로 알아차려야 한다. 사실은 이렇다. EA 운동은 적어도 그 실천에서는 굉장히 공리주의적이다. 실제로 그 이름이 결정되기 전에, Ord 를 포함한 EA 운동의 초기 멤버들은 EA 운동을 효과적 공리주의 모임으로 부를 것을 심각하게 고민했다.

그러니까, 공리주의는 비실체적이고 공정한 우주적 관점에서 합산된 세계의 본질적인 가치의 총량을 최대화하는 것을 유일한 도덕적 의무로 삼는 도덕 이론이다(?, This being said, utilitarianism is an ethical theory that specifies our sole moral obligation as being to maximise the total amount of intrinsic value in the world, as tallied up from a disembodied, impartial, cosmic vantage point called the point of view of the Universe). 이 관점에서, 시간과 공간에 따라 사람들에게 (공리적 쾌락주의자 입장에서는 즐거움과 같은) 가치가 어떻게 분배되는지는 중요하지 않다. 중요한 것은 가치의 총합이다. 예를 들어서, 1의 가치를 갖는 1조 명의 사람이 있다고 생각해 보자. 이는 간신히 살 가치가 있는 것이다. 이 경우에 가치의 총합은 1조다. 이제 10억 명이 999의 가치를 가지고 있는 다른 세계를 생각해 보자. 9990억은 1조보다 작기 때문에, 살아야 할 가치가 겨우 있는 사람들로 가득찬 첫 번째 세계가 두 번째 세계보다 도덕적으로 낫다. 따라서, 만약 공리주의자에게 이 두 세계 중 한 세계를 선택하도록 강요한다면, 그는 첫 번째 세계를 꼽을 것이다. (이것은 당혹스러운 결론으로 불리는데, Ord, MacAskill 그리고 Greaves 같은 장기주의자들은 현재 이 문제가 크게 심각하지 않다고 주장한다. 이들에게는 첫 번째 세계가 정말로 좋을 수 있다!)

## Beckstead는 가난한 나라의 사람들보다 부유한 나라 사람들의 삶을 더 우선시 해야 한다고 주장했다.

위 인용구의 기저에 깔린 논리는 사람은 목적을 위한 수단 이상의 것이 아니라는 생각에 기반한다. 우리는 그 자체로 중요하지 않고, 우리 자신만의 고유한 가치도 없다. 대신에 인간은 가치를 담는 그릇으로 이해되고, 따라서 우리는 가치를 지니고 있는 경우에만 중요하고, 그래서 빅뱅과 열 평행 사이의 우주에서 전체적인 가치의 총량에 기여한다. 공리주의자들은 우리에게 가치를 최대화 해야 한다고 말한다. 이는 윤리적으로 말해서 (즐거움이나) 가치의 총량이 양(+)인 상태로 존재하는 (가치의 그릇인) 사람이 더 많아질수록 우주가 더 좋아진다는 결론에 이른다. 한 마디로, 가치가 사람을 좋게 만들기 위해 존재하는 것이 아니라 사람은 가치를 최대화하기 위해 존재한다.

이것이 왜 장기주의자들이, 우주를 식민지화하고 별 주위에 가상 환경에서 양의 가치를 갖고 살아가는 엄청나게 많은 수의 사람이 있는 거대한 컴퓨터 시뮬레이션을 만들 수 있다고 가정한 뒤에, 미래에 존재할 수 있는 사람들이 얼마나 되는지 계산하는 것에 집착하는 이유다. 내가 이미 언급했지만, Bostrom은 다양한 디지털 인간을 포함해서 10의 54승의 미래 인구를 추정했다. 그의 베스트셀러인 <<슈퍼인텔리전스>>에서 이 숫자를 10의 58승으로 더 높아지기는 했지만, 이 사람들은 거의 대부분은 가상 환경에서 타인과 상호작용하며 부유하고 행복하게 살 것이라고 예상했다. 유사하게 Greaves와 MacAskill은 단지 우리 은하 안에서도

10의 45승 정도 되는 의식적 존재가 컴퓨터 시뮬레이션 안에 존재할 수 있다고 추정하면서 이 가능성에 대해 흥분하고 있다.

이것이 우리 미래의 빛 원뿔 도처에 놓여있는 거대한 컴퓨터 시뮬레이션 안에 기술적으로 강화된 엄청나게 많은 디지털 포스트 휴먼으로 구성된 광대하고 영광스러운 잠재력이다. Häggström의 시나리오에서 이 목표를 위해 장기주의 정치인들은 독일을 멸망시킬 수 있다. 이 목적을 위해 우리 자원을 세계의 빈곤 문제를 해결하는데 낭비해서는 안 된다. 이 목적을 위해, 세계적 감시 시스템을 구축하는 것을 고려해야 하고, 선제적 전쟁을 선택 사항으로 여겨야 하고, 그리고 (대부분 북반구 사람에 의해 유발된) 기후 변화의 절망적 영향에서 남반구 사람을 구하는 일 보다는 인공지능 기계에 더 집중해야 한다. 사실 Beckstead는 심지어, 이 목적을 달성하기 위해서, 우리가 가난한 나라의 사람들보다 부유한 나라 사람들의 삶을 더 우선시 해야 한다고 주장했다. 왜냐하면 압도적으로 중요하고 부유한 사람들이 가난한 사람보다 장기적 미래에 더 영향을 미칠 것이기 때문이다. Ord가 장기주의 저술에 가장 중요한 기여 중 하나라고 열정적으로 찬사를 보낸 Beckstead의 박사 논문에서 한 구절을 인용하겠다.

가난한 나라에서 생명을 구하는 것은 부자 나라에서 생명을 구하고 개선하는 것보다 현저하게 작은 파급 효과를 가질 수 있다. 왜냐하면, 부유한 나라가 더욱 혁신적이고 그곳의 노동자가 경제적으로 더 생산성이 있기 때문이다. [결과적으로] 나에게서는 다른 상황이 같다면, 가난한 나라의 생명을 구하는 것보다 부유한 나라의 생명을 구하는 것이 더 많이 중요하다고 말하는 것이 더 그럴듯하다.

---

이것은 빙산의 일각에 불과하다. 우리의 잠재력이라는 개념이 기술의 발전과 새로운 위험의 발생에 미치는 영향을 생각해보자. 우리 잠재력을 실현하는 것이 인류의 궁극의 도덕적 목표이기 때문에, 그리고 오늘날보다 더 진보된 기술이 없으면 우리 후손이 포스트 휴먼이 될 수 없고, 우주 식민지를 만들 수 없고 그리고 컴퓨터 시뮬레이션에 들어가는 10의 58승의 인구를 탄생시킬 수 없기 때문에, 기술 개발에 실패하는 것은 그 자체로 존재적 재앙이다. 이는 Bostrom이 안정기라고 부르는 실패의 한 방법이다. (Ramsey가 당구와 음주로 자신의 재능을 무시한 것과 비교해 보라.) 실제로 Bostrom은 이 생각을 존재적 위험에 대한 그의 표준적 정의에 중심에 놓았다. 이 위험은, 인류가 기술적 성숙을 실현 가능하게 성취할 수 있는 최댓값에 가깝게 본성을 조절하고, 일정 수준의 경제적 생산성을 획득한 기술적 성숙 상태에 이르거나 유지하는 것을 방해하는, 미래의 사건을 의미한다. 기술적 성숙이 중요한 이유는 절대적인 물리적 제한까지 경제적 생산성을 올리고 본성을 조절하는 것이 미래의 빛 원뿔 안에서 가치의 양을 최대화하기 위한 표면적인 요구사항이기 때문이다.

인류가 어떻게 현재 기후 위기와 환경 위기에 스스로 빠지게 되었는지 잠시 생각해보자. 화석 연료의 추출과 연소, 대규모 환경 파괴와 종의 절멸 이면에는 자연을 통제하고, 정복하고, 착취하고, 약탈하고, 부수고, 변형하고, 재구성하고 조작해야 할 대상이라고 보는 관념이 있다. 기술 이론가 Langdon Winner 는 *Autonomous Technology* 에서, Francis Bacon 시대부터, 기술에 대한 우리의 관점은 (절대적 지배의 스타일, 독재, 노예에 대한 주인의 일방적 통제같이) 권력이 사용되는 방식 대한 단일 개념과 불가분하게 엮여 있었다고 말한다. 그는 다음을 덧붙인다.

자연의 모든 것을 정복하고, 차지하고, 지배하는 인간의 권리에 대해 유보 조항은 찾기 어렵다. 이것이 인간의 힘과 영광이다. 다른 상황에서는 다소 천박하고 비열한 의도로 보일 수 있는 것이 여기서는 가장 명예로운 미덕이 된다. 인간이 자연을 조작하는 것이 적절하다고 스스로 생각하기 때문에 자연은 보편적인 먹잇감이다.

이것은 정확히 Bostrom 의 존재적 위험과 그에 따른 규범적 미래학에서 찾을 수 있는 것이다. 자연, 우주 전체, 우리의 우주적 자산은 약탈을 위해 존재한다. 그리고 이것들은 Bostrom 이 "Astronomical Waste"에서 말하는 것처럼, 거대한 컴퓨터 시뮬레이션에서 가치 있는 삶을 사는 감각적 존재와 같은, 가치 구조로 조작되고 변형되고 전환된다. 그러나 이런 베이컨주의적이고 자본주의적인 관점은 생물권의 넓은 부분과, 전 세계 원주민 공동체, 심지어 서구 기술 문명 그 자체의 파괴를 당장 위협하는 전례 없는 환경 재앙을 불러일으키는 가장 근본적인 원인이다. 다른 장기주의자들이, Bostrom 처럼 명확하지는 않지만, 공리주의가 인간을 보는 방식으로, 그러니까 추상적이고 비인격적인 목적을 위한 수단 그 이상이 아닌 것으로, 자연을 보는 경향은 분명하다. 예를 들어서, MacAskill 과 한 동료는 사실상 장기주의인 EA 운동은, 선한 일을 한다는 잠정적 목표가, 다른 목적 없이 생물다양성을 보존하거나 자연의 아름다움을 보존하는 것이 아니라, 행복을 증진하는 것에 초점을 맞춘다는 점에서, 잠정적 **복지주의**라고 말한다.

## 이런 이유로 모든 문제는 기술이 과해서라기보다는 부족해서 생긴다.

걱정스러운 것은, 오늘날 인류 멸종의 가장 두드러진 원인은 바로 이 기술 때문이라는 합의된 사실에도 불구하고, 장기주의자들은 언제나 더 강력한 기술을 개발해야만 한다고 주장한다는 것이다. Ord 에 따르면, 인류를 지키기 위한 진지한 노력이 없다면, 이번 세기에 인류가 절멸할 위험이 더 높아질 것이고, 기술 발전이

계속되는 세기마다 그 위험은 증가할 것이라고 믿을 강력한 이유가 있다. 이와 유사하게 2012 년에 Bostrom 은 다음과 같이 말했다.

가까운 미래에 존재적 위험의 대부분은 인간이 만들어낸, 즉 인간 활동으로 발생하는 존재적 위험이다. 특히, 가장 큰 존재적 위험의 대부분은, 우리의 생물학적 측면이나 외부 세계를 조작할 수 있는 능력을 근본적으로 증대시킬, 미래에 일어날 수 있는 기술적 도약과 관련된 것으로 보인다. 우리 능력이 증대됨에 따라서, 의도했든 아니든, 긍정적이든 부정적이든, 잠재된 결과의 규모도 커질 것이다.

이 관점에서 앞으로 나아갈 수 있는 유일한 길은 기술 발전뿐이다. 이 길이 미래로 가는 가장 위험한 길일지라도 말이다. 그러나 이것이 얼마나 의미가 있을까? 확실히, 생존 확률을 최대화하는 것을 원한다면, 우리는 민간과 군용으로 모두 쓰일 수 있는 새로운 기술의 발전에 반대해야만 한다. 만약 (역사에서 명확히 보여주고 기술 프로젝트에서 확인할 수 있듯이) 기술 발전이 위험을 늘리는 것과 동일하다면, 아마도 존재적 안전 상태를 확보하는 사실상 유일한 길은 더 큰 기술 혁신을 중지하거나 느리게 만드는 것이다.

그러나 장기주의자들은 이 난제에 대한 답으로 가치 중립 이론을 가지고 있다. 가치 중립 이론은 기술이 도덕적으로 중립인 대상, 즉 단순히 도구에 불과하다는 것이다. 이 생각은 "총은 사람을 죽이지 않는다. 사람은 사람이 죽인다"라는 전미 총기 협회의 슬로건에 잘 포장되어 있다. 이 슬로건은 기술의 결과는, 선하든 악하든, 이익이 되든 해롭든, 인공물이 아닌 사용자에게 의해 결정된다는 의미를 갖고 있다. 2002 년에 Bostrom 은 "우리는 문명과 기술이 거대한 존재적 위험을 불러온다고 비난해서는 안 된다"라고 말하면서 "왜냐하면 존재적 위험을 정의하는 방식에 따르면, 기술 문명 발전에 실패한 것은 이미 우리가 존재적 재앙의 희생자로 전락한 것을 의미할 수 있다"라고 덧붙인다.

이와 유사하게 Ord 는 Carl Sagan 의 *Pale Blue Dot* 을 인용하기 전에 "문제는 지혜의 부족이지 과도한 기술이 아니"라고 주장한다. Sagan 은 "우리가 처한 많은 문제들은 과학과 기술에서 발생하는 것이 아니라 더 근원적으로 우리가 가진 힘에 비해 충분한 지혜를 갖고 있지 않기 때문이다"라고 말했다. 그러니까, 더 영리하고 더 지혜롭고 그리고 더 윤리적이지 못하게 된 것이 우리 책임이라는 말이다. 이런 결핍의 묶음에 대해 많은 장기주의자들은 다소 왜곡된 논리로 우리의 인지 체계와 도덕적 기질을 기술적으로 재설계함으로써 교정할 수 있다는 믿음을 가지고 있다. 이런 이유로 모든 것은 공학적인 문제이고 따라서 모든 문제는 기술이 과해서라기보다는 부족해서 발생한다.

이제 우리는 장기주의가 얼마나 자멸적인지 알 수 있다. 우리의 장기적 잠재력 실현에 대한 장기주의의 광신적 강조는 사람들이, 예를 들어, 비 존재적 기후 위기를 등한시하게 만들고, 부자를 빈자보다 중요하게 여기게 하고, 심지어 어쩌면 어 더 거대한 우주적 선을 위해 선제적 폭력과 극악한 행위를 정당화하게 이끈다. 그뿐만 아니라 장기주의는, 베이컨주의, 자본주의 그리고 기술에 대한 가치 중립적 태도 같은, 인류를 파멸의 벼랑 끝으로 밀어내는 성향을 가지고 있다. 장기주의는 우리에게 경제적 생산성, 자연에 대한 통제, 우주에서 우리의 존재, 미래에 존재할 (시뮬레이션에서 살아갈) 사람의 수, 비인격적인 가치의 총합 등등을 최대화해야 한다고 말한다. 그러나 최대화를 위해서는 지속적으로 뛰어난 (그리고 위험한) 기술을 개발해야만 한다. 여기서 실패하는 것은 그 자체로 존재적 재앙이다. 걱정할 필요는 없다. 그래도 기술이 우리가 악화되는 상황에 책임이 없기 때문에 위험 대부분은 기술에서 직접적으로 기인한다는 사실이 기술 개발을 멈춰야 한다는 이유가 되지 않는다. 오히려 문제는 우리에게 있다. 이는 우리 자신을 인지적이고 도덕적으로 강화된 포스트 휴먼으로 변화시킬 수 있는 기술을 개발해야만 하는 것을 의미한다.

이것은 재앙을 불러일으키는 방법으로 보인다. 지혜롭고 책임감 있는 포스트 휴먼이라는 새로운 인종을 만드는 일은 그럴듯하지 않다. 만약 진보된 기술이 지금과 같은 속도로 계속 개발된다면, 지구적 규모의 재앙은 만약의 문제가 아니라 언제의 문제가 될 것은 거의 확실하다. 물론, 대략 십억 년 안에 태양에 의해 지구가 살균되기 전에 탈출하기를 원하면 발전된 기술이 필요할 수 있다. 그러나 장기주의자들이 놓치고 있는 중요한 사실은 **기술이 장기주의 자들이 생각하는 먼 미래가 오기 전에 우리를 절멸시키는 원인이 될 가능성이 우리를 구원할 가능성보다 훨씬 높다는 것이다.** 만약 당신이 나와 같이 지속적인 생존과 인류의 번역을 중요하게 여긴다면, 당신은 장기주의의 이데올로기를 거부하고 장기적 시간을 생각해야 한다. 장기주의는 위험하고 결점이 있을 뿐만 아니라 지금 우리를 위협하는 위험에 기여하고, 이 위험을 강화한다.