

일론 머스크 테슬라모터스 대표, 스투어트 러셀 UC 버클리대 컴퓨터공학과 교수, 바트 셀맨 코넬대 컴퓨터공학과 교수, 미래학자 레이 커즈와일, 철학자 데이빗 차머스 뉴욕대 교수, 닉 보스트롬 옥스퍼드대 인류미래연구소장, 데미스 하사비스 딥마인드 대표, 샘 해리스, 그리고 스카이프 창업자인 얀 탈린, 맥스 테그마크(사회자)

맥스 테그마크: 지금부터 질문을 드릴게요. 대답은 '예', '아니오' 또는 '난해하다'로 대답해주세요. 아시겠죠? 여기서부터 시작하겠습니다. 형태를 무관하고, 과연 초지능(superintelligence)은 가능할까요? '예', '아니오' 또는 '난해하다'로 답해주세요.

바트 셀맨, 레이 커즈와일, 데이빗 차머스, 닉 보스트롬, 데미스 하사비스, 샘 해리스: 예.

스튜어트 러셀: 확신합니다.

일론 머스크: 아니오.

모두: (웃음)

맥스 테그마크: 불일치한 의견을 못 찾아서 정말 아쉽네요. 좀 더 분발합시다. 가능성이 있다고 해서 실제로 일어난다고 단정 지을 수는 없겠죠. 그렇다면 물리학의 법칙에 따라 초지능(superintelligence)은 가능할 것인가에 대한 질문을 드리기 전에, 실제로 일어날 것이라고 생각하십니까?

얀 탈린: 조금 난해 하지만 예.

샘 해리스: 예, 만약 일어나지 않았다면 그것을 가로 막는 끔찍한 사건이 일어났을 겁니다.

데미스 하사비스 : 예.

닉 보스트롬: 아마도요.

스튜어트 러셀, 바트 셀맨, 레이 커즈와일, 데이빗 차머스 : 예.

일론 머스크: 아니오. (※일론 머스크의 장난섞인 답변, 그는 실제로는 가능하다는 답변을 보였습니다)

모두: (웃음)

맥스 테그마크: 이번에도 불일치한 의견을 찾지 못했네요. 조금 더 분발 해야겠습니다. 좋습니다. 모두 초지능은 실제로 일어날 것이라고 생각 하시는군요. 하지만 실제로 일어나기를 원하십니까? '예', '아니오' 또는 '난해하다'로 대답해주세요.

얀 탈린: 난해합니다. 하지만, 원하는 쪽에 더 가깝네요.

샘 해리스: 난해합니다.

닉 보스트롬, 데미스 하사비스: 예.

데이빗 차머스: 매우 난해합니다.

레이 커즈와일: 예.

바트 셀맨: 난해합니다.

스튜어트 러셀: 매우 난해합니다.

일론 머스크: 음.. 잘 모르겠네요. 종류에 따라 다를 거 같아요.

맥스 테그마크: 좋습니다. 차차 흥미로워지고 있네요. 과연 언제 일어날까요? 오전 패널에서 인간 수준의 인공지능은 언제 도달할 것인가에 대해 흥미로운 토론을 진행했죠. 따라서, 언제 시작할 것인가에 대한 질문을 다시 꺼낼 필요는 없을 것 같네요. 방금 이야기한 것과 매우 관련된 질문을 해봅시다. 빠른 진화의 과정에서 시간에 대한 예측 전망은 상이해 지기 마련입니다. 이 부분에 대해서는 의견의 차이가 생길 가능성이 있겠네요. 어떤 사람들은 '붐!' 하고 며칠, 몇 시간 아니면 더 짧은 시간 내에 일어나는 시나리오를 구상하고 있습니다. 반면 다른 사람들은 분명 일어날 것이라고 예상하지만 수 천 년 또는 수 십 년이 걸릴 것이라고 생각하고 있습니다. 여기 계신 분들이 생각하시는 개략적인 타임 스케일은 언제쯤 입니까? 안?

얀 탈린: 지금이요? 아니면 인간 수준부터?

맥스 테그마크: 아니요 그러니까 인간 수준의 인공지능을 도달한 시점 또는 그 이상 아니면 그 이하도 좋습니다. 모든 것이 본격적으로 시작하는 시점을 말하는 겁니다. 예를 들어, 어떤 폭발이 일어나기 전에 물리학자들이 예측하는 타임 스케일이 있듯이, 예상 타임 스케일은 어느 정도 입니까? 몇 초, 몇 년 아니면 수천년?

얀 탈린: 제 생각엔 몇 년이라고 생각합니다만, 단기간이라는 가정을 가지고 대비해야 합니다.

샘 해리스: 사실 저는 제 직감을 믿지 않아요. 하지만 제 직감은 수년이라고 예상합니다. 동시에 저는 인간 수준의 인공지능은 신기루라고 생각해요. 언젠간 갑자기 인간보다 나아질 것입니다. 그러나 빠른 시일 내에 초지능의 시대가 완벽하게 도래할 것인지는 모르겠습니다.

데미스 하사비스: 부분적으로 인간 수준 인공지능을 가능케 하는 것은 설계에 의존한다고 생각합니다.

인공지능은 지금 우리가 연구하고 있는 신경과학에서 계승됐습니다. 그리고 이런 인공지능의 지능을 향상시키기 위해선 학습과 경험이 필요합니다. 몇 년 또는 수십 년이 걸릴 수도 있겠네요.

닉 보스트롬: 몇 년? 근데 훨씬 단축 될 수도 있어요. 그러나 반대로, 더 오래 걸린다 해도 놀랍지 않을 겁니다.

데이빗 차머스: 어쩌면 며칠 또는 그 이하가 될 수 있습니다.

레이 커즈와일: 인공지능의 발전이 있을 때마다, 체스, 바둑, 자율주행차량처럼, 우리는 '에이, 그건 진짜 인공지능이 아니죠'하며 무시해버리는 경향이 있어요. 인공지능은 아시다시피 아직 도전해보지 않은 분야입니다. 그리고 이런 경향은 범용 인공지능(Artificial General Intelligence, 이하 AGI)에 실제로 도달해도 계속 될 거예요. 많은 논란이 있을 것입니다. 논쟁이 진정 될 즈음에 우리는 인공지능이 몇 년 동안 주변에 존재했다는 것을 깨닫게 되겠죠.

바트 셸맨: 저는 여러 분야에서 인간의 능력을 뛰어 넘을 것이라고 생각해요. 하지만 모든 분야가 동시에 그렇지는 않겠죠. 불균등한 과정이 될 거예요. 어느 분야는 매우 빨리 훨씬 더 발전 할 것이고 어느 분야는 훨씬 오래 걸릴 겁니다.

스튜어트 러셀: 바트의 의견에 동의합니다.

일론 머스크: 가장 똑똑하고 독창적인 인간만큼 인공지능을 발달하는 한계점에 도달한다면.. 모든 인류의 지식을 합친 것보다 더 똑똑한 인공지능을 만드는 것은 시간 문제겠죠.

맥스 테그마크: 여기서 우리는 꽤 흥미로운 답변을 보았습니다. 개인적으로 방금 드린 질문을 매우 흥미롭게 생각하고 있습니다. 왜냐하면 여기 계신 분들이 인공지능의 시기를 주제로 한 흥미로운 논문을 발표해왔기 때문이죠. 맞습니다. 안이 적절한 비유를 해주셨어요. 인공지능 기술의 진보에 맞춰 사회가 적응할 수 있어야 합니다. 반면에 사회가 대응할 수 있는 시간보다 훨씬 빨리 발달되는 경우 통제가 훨씬 더 힘들어지고, 통제 방안을 사전에 미리 구축했기를 바랄 수 밖에 없겠죠. 여기 계신 분들도 같은 생각을 하고 있었지만 제가 궁금한 것은 지능의 성장에 대한 궁금증입니다. 이상적인 인공지능은 우리가 모두 원하는 바이고 이에 맞춰 사회가 적응할 수 있어야 합니다. 그렇게 하기 위해선 지능의 성장을 늦추는 것도 생각해봐야 합니다. 혹시 패널 분들 중에서 지능의 성장을 늦추는 것에 찬성하시는 분 있습니까? 만약 그렇다 하신 분들을 위해 좀 더 어려운 질문을 드리겠습니다. 천천히 적응할 수 있도록 지능의 성장을 늦출 수 있는 방법은 어떤 것이 있습니까? 모두에게 드리는 질문은 아닙니다. 답변 해주실 분 계십니까?

얀 탈린: 푸에르토리코에서 리치 서튼과 나눈 대화가 생각나네요. 의견이 충돌이 많이 있었죠. 하지만 빠르게 진행되는 것보다 느리게 진행되어야 한다는 점에서는 둘 다 동의 했습니다.

맥스 테그마크: 혹시 구체적인 방법에 대한 의견이 있으신 분?

레이 커즈와일: 제가 발표 자료에서 제안한 전략은 어느 정도 능담이 섞여 있었어요. 하지만 진지했던 부분도 있었습니다. 저는 이 컨퍼런스의 취지가 훌륭하다고 생각합니다. 기술자로서 기술의 안전과 유익을 보존하기

위해 모든 것을 시도해야 한다고 생각합니다. 물론, 자율주행차량 같은 응용 기술을 개발할 때 해결이 필요한 윤리적 문제가 많이 발생합니다. 그러나 이런 문제에 대한 해답을 기술로만 해결할 수 있다고 생각하지 않습니다. 우리가 모든 과정을 오차 없이 진행했고, 가장 안전하고 유익한 인공지능을 만들었다고 상상해봅시다. 모든 것이 완벽하지만 우리의 정치 체제는 악랄합니다. 전체주의적인 시스템이 될 때까지 아무것도 하지 않았어요. 일부 정치가 그렇든 전체가 그렇든, 이런 식으로는 잘 풀리지 않을 거예요. 정치와 사회 정책은 우리가 원하는 가치를 사회가 반영할 수 있도록 더 개선되어야 합니다. 우리가 만드는 것은 인간 인공지능 이에요. 정의상 인간 인공지능은 인간 수준이라는 뜻입니다. 따라서 똑같은 인간이라고 간주해야 해요. 도덕적인 인간을 만드는 방법은 인간 수준의 윤리를 갖춘 인공지능을 만드는 방법과 같습니다.

맥스 테그마크: 인간 수준의 인공지능에 가까워 지고, 사전에 구비해야 할 가장 좋은 대응책은 인류 즉, 인간 사회 모두가 힘을 합쳐 함께 노력하는 것. 그리고 최소한 현재보다 최대한 많은 근거를 가지고 행동하는 것이 중요하다는 말씀이죠? 맞습니까?

레이 커즈와일: 정확합니다.

맥스 테그마크: 닉? 아, 그리고 확인 차 다시 얘기해드립니다. 제가 인공지능의 성장을 늦추는 것에 대한 질문을 드렸을 때 당장 인공지능에 대한 연구를 늦추자는 의미가 아닙니다. 인간 수준의 인공지능에 가까워지고 진행 속도에 불이 붙는다는 가정하에, 이 부분을 느리게 하는 방법은 어떤 것이 있을까요?

닉 보스트롬: 한 가지 방법은 현재 진행 속도를 촉진 시켜 하드웨어가 덜 발전되었을 때 그 지점에 먼저 도달하는 겁니다. 그러면 불필요한 하드웨어를 줄일 수 있겠죠. 하지만 현재 인공지능 연구의 진행 속도는 아주 큰 충격이 발생하지 않는 이상 느려지기 쉽지 않습니다. 그렇기 때문에 제 발표 차례에서 말씀 드렸듯이, 가장 유력한 옵션은 우리가 예상하고 있는 한계점에 가장 먼저 도착한 사람이, 적어도 몇 달 정도는, 뒤따라오는 사람들을 천천히 이끌 수 있는 역량이 있어야 합니다.

스튜어트 러셀: 제 생각에 한 가지는 할 수 있어요. 아직은 대부분이 검증의 영역에 있지만, 새로운 하드웨어를 만드는 등의 도전을 감행할 수 있어요. 이를 통해 자원의 보다 효과적인 활용이 가능할 것입니다. 이렇게 되면 저는 몇 년 동안 가만히 앉아서 다음 단계에 무엇을 할 것인가에 대해 열심히 고민하고 있을 겁니다. 생각만해도 아주 행복하네요.

레이 커즈와일: 하지만 소프트웨어를 복제하는 것은 너무 쉬운 일이에요. 소프트웨어는 언제나 그랬듯이 자기 복제 기능을 갖추고 있어요. 그 기능을 중지 할 수 있는 방법이 존재하긴 하는지 모르겠네요.

데이빗 차머스: 시간을 벌어서 진행 속도를 늦출 수 있으면 정말 좋겠죠. 하지만 최초의 초지능 개발자나 퍼스트 무버에게 주어질 이익을 생각해보면 늦추는 것은 불가능에 가깝다고 생각합니다. 제가 생각하는 유일한 시나리오는 단 한 명의 퍼스트 무버가 잠시 진행을 중단하는 겁니다. 생각했던 시나리오가 현실이 되면 인공지능을 질서 있게 통제할 수 있는 사회를 볼 수도 있겠네요.

맥스 테그마크: 네 데미스, 우리의 동료인 웨인 레그도 이런 컨퍼런스처럼 여러 분야가 협력해서 주도자들 사이의 신뢰를 쌓고 서로 개방한다면 도움이 될 거라고 언급 했었죠. 혹시 다른 의견...

데미스 하사비스: 네 저도 다양한 시나리오를 포함한 여러 방면에서 고민을 많이 해요. 개인적으로 시간이 충분하다면 이러한 문제들은 인간의 창의력으로 해결할 수 있다고 아주 강하게 믿습니다. 통제 및 기타 문제들은 어렵지만 해결할 수 있다고 생각해요. 진짜 문제는 나중에 실제로 진행 속도를 늦춰야 하는 시점에서 서로 조율하다가 시간을 허비하는 거죠. 만약 한 주제에서 스템어가 5년 동안 고민하는 동안 어떤 사람은 찬성하고 다른 팀은 반대하고, 이런 사태가 걱정되는 겁니다. 이렇게 사람들끼리 조율하는 것은 어렵게 느껴질 수 있어요. 하지만 첫 번째 단계로 인공지능 협회를 설립하거나 가장 능력이 뛰어난 팀들이 어떤 조약이나 안전성 절차에 합의하는 등, 이런 것들이 도움이 많이 되겠죠. 이런 과정도 몇 년은 걸릴 거예요. 그리고 진행 속도를 '늦춘다'와 '빠르게 진행한다'에 대한 논쟁에 대해 한가지 말씀드리고 싶은 게

있습니다. 인공지능 연구가 50 년동안 일시적으로 중지되고, 하드웨어는 지금처럼 계속 증가한다고 상상해보세요. 이견 닉의 의견과 비슷하지만, 이런 경우 씨드 인공지능 또는 자기 개선 인공지능 등 인공지능에 대한 다양한 시도가 이루어질 수 있었던 시간에 대량의 하드웨어 과부하가 발생합니다. 결국엔 전문가가 아닌 일반인이 직접 개발할 수도 있어요. 이런 상황이 닥치면 합의점을 찾는 과정이 더 어려워질 뿐입니다. 그러므로 어려운 국면에 도달했을 때 진행 속도에 대한 합의가 이루어져야 할 필요가 있습니다. 그 단계는 아마 더 큰 규모의 연구진이 필요할 거고 시야를 넓혀야 해요. 주요 리더가 누군지 그들이 어떤 연구를 하고 있는지 잘 파악하고 있어야 합니다. 이와 상반되는 시나리오로 50 년 또는 100 년의 기간 동안 전문가가 아닌 일반인이 씨드 인공지능과 같은 것을 발명하는 불상사가 발생할 수도 있습니다.

맥스 테그마크: 알겠습니다. 바트, 이 주제에 대한 마지막 의견을 주시죠.

바트 셸맨 : 네. 저는 이 과정이 매우 불규칙하게 이루어질 것이라고 생각해요. 어느 때는 훨씬 발전하고 어느 때는 아주 느리게 진행될 겁니다. 그리고 몇몇 사람들은 기술과 기계를 이용해서 가짜 영상 즉, 증명되지 않은 내용이 담긴 영상을 퍼트리죠. 이런 영상이 돌아다닐 때 한 사회로서, 사람들이 좀 더 민감하게 받아들였으면 하는 바람입니다. 완전한 범용 인공지능이 나오기 전에 이런 문제를 좀 더 심각하게 받아들였으면 좋겠어요.

레이 커즈와일 : 그건 인공지능을 활용해서 판단하면 되겠네요.

맥스 테그마크: ‘걱정’이란 단어를 거론하셨었네요. 그리고 닉, 발표 슬라이드에서 ‘파멸’이란 단어를 3 번 쓰셨고요. 아마존에서 별 1 개를 받은 이유를 알겠네요. 게다가 빨간색이었어요. 저는 이런 걱정과 우려에 대한 토론이 중요한 만큼 현실적인 희망과 환상적인 장점에 대해 얘기하는 것도 중요하다고 생각합니다. 그리고 그런 장점에 대한 토론을 이 자리에서 나누었으면 좋겠습니다. 일단 지금부터는 걱정과 우려를 빠르게 짚고 넘어간 뒤 다시 긍정적인 이야기로 돌아가봅시다. 여러분들이 생각하고 있는 반드시 짚고 넘어가야 할 어려운 과제 중 한가지와 그것을 최소화 시키는 최선의 방법에 대해 한 명씩 빠르게 얘기해봤으면 좋겠네요. 양부터 시작 해보실래요?

얀 탈린: 어떤 것을 최소화 해요?

맥스 테그마크: 생각하고 계시는 리스크 중 한가지에 대해 얘기해주세요. 그리고 그것을 최소화시킬 수 있는 체계적인 방법에 대해 얘기해주세요.

얀 탈린: 개인적으로 서로 인공지능 기술과 정보를 얻기 위해 경쟁하는 태도에 대해 긍정적입니다. Open AI 와 DeepMind 가 서로 협조하는 것도 정말 좋은 현상이죠 그리고 이것은 단지 더 큰 그림의 일부라고 생각합니다. 최근에 저는 아시아에서 주로 활동했습니다. 다양한곳에서 더 많은 사람들을 끌어 들이고 영입하기 위해서였죠. 왜냐하면 지금까지는 주로 영미 지역 위주로 진행되었기 때문입니다. 그렇기 때문에 다양한 곳에서 사람들을 끌어들이는 것은 꼭 해야 할 일이고 제 자신 또한 노력할 겁니다.

샘 해리스: 이 분야의 전문 지식이 없는 사람으로서 짚고 넘어야 할 리스크는 안전성에 대한 문제를 감정적으로 심각하게 받아들이지 않고 있다는 것입니다. 그리고 전문가들 또한 심각하게 받아들이기 힘든 것도 사실이죠. 외부에서 자신을 전문가라고 자칭하는 사람들이 통제에 대한 문제를 전혀 문제가 되지 않는다고 생각하는 이들이 얼마나 많은지 모르실 겁니다. 이런 사람들을 만나면 정말 놀라워요. 그래서 생각할만한 가치도 없죠. 제가 우려하고 있는 것은 시간의 축이 중요하다는 착각입니다. 50 년 또는 100 년의 시간이 흐르면 통제가 가능한, 안전한 인공지능을 만들기 충분하다고 암묵적으로 가정하고 있는 거죠. 시간의 축이 50 년 또는 100 년이라는 가정을 우리가 인지하고 있기 때문에 불안한 마음을 못 느끼는 겁니다.

두 번째로, 사람들은 인공지능을 본질적인 가치라고 착각하고 있어요. 당연히 사람들은 더 많이 원하고 있고 쉽게 착각에 빠질 수 있습니다. 암 치료법이 존재한다는 것을 지금은 발견했지만, 치료법에 대한 정보가 없죠. 분통이 터지는 현실입니다. 미래에는 어떤 실험을 진행해야 하는지, 가지고 있는 데이터를 어떻게 활용한 것인지 등등 암을 치료하기 위한 더 많은 정보를 발견하겠죠. 따라서 깊이 생각했을 때, 충분한

정보를 가지고 있지 않다는 것은 몹시 괴로운 일입니다. 아까 물어보신 질문으로 다시 돌아가서, 진행 속도를 촉진시키면 인공지능은 본질적인 가치가 되 버릴 겁니다. 인공지능에 관련된 정치적 문제와 기타 문제를 이미 해결했기 때문에 혼란이 발생하지 않을 것이라는 착각 속에 빠지겠죠. 저는 사람들이 이러한 문제들을 가볍게 여기고 있다고 생각합니다.

맥스 테그마크: 샘, 제가 들리는 바로는 당신이 웨인 레그의 의견에 강하게 동의하고 있는 것 같네요. 몇몇 사람들은 AGI를 터무니 없다고 생각하고 AGI의 가능성을 무시하죠. 그는 이런 사람들의 생각에 변화를 준다면 신속하게 연구를 진행할 수 있고, 여러가지 문제에 대한 해결책과 답을 찾을 수 있다고 했죠. 그렇다면 만약 저와 단 둘이 엘리베이터 안에서 30초의 시간이 있다고 가정하고 당신한테 '초 인간 지능의 아이디어는 말도 안되고 우스꽝스럽다'라고 말했다면 어떻게 저를 설득시킬 건가요?

샘 해리스: 음.. 이런 것들을 조금 더 진지하게 받아들이기 위해 몇가지 추측을 해야합니다. 다시 말하지만, 감정적인 부분은 별개입니다. 지능을 단지 물리적 시스템으로 정보를 추출한 생산물로 가정하면 위협적이지 않을 겁니다. 그리고 레이가 얘기했던 것처럼, 만약 신이 완벽하게 친화적이고 모범적인 인공지능을 인간에게 제공했다고 해도, 현재 정치적 및 경제적 시스템을 고려했을 때 큰 혼란을 가져올 겁니다. 현재 우리는 윤리적 그리고 정치적 의지가 부족해서 인공지능의 능력이 평등하게 분배될 가능성이 없습니다. 실리콘 밸리와 중국 또는 이란도 정치적으로 분리되어 있죠. 닉이 발행한 책에서 인용하자면, 인공지능의 문제점을 해결했음에도 불구하고 리스크는 존재합니다. 빠른 시일 내에 해결책을 찾는게 중요합니다.

맥스 테그마크: 감사합니다 샘. 데미스, 당신이 우려되는 과제와 집중적으로 조치할 것에 대해 설명해주시겠습니까?

데미스 하사비스: 방금 까지 나왔던 의견에 동의합니다. 그리고 조율 문제가 중요하다고 생각해요. 경쟁 속에서 오로지 목적을 달성하기 위해 원칙을 무시하는 행위가 발생하는 것을 방지하기 위해 말이죠. 안전성은 인공지능의 실질적인 능력에 부합하지 않기 때문에 무시 당하기 쉽다고 생각합니다. 세상이 주목해야 할 이슈 중 하나입니다. 그리고 현재 정치 체제를 감안하면 쉽지 않은 문제가 될 겁니다. 사회에서 활동하는 인공지능은 어떤 모습일 것인가? 개수는 어느 정도일 것인가? 그들의 목표는 누가 정할 것인가? 이러한 정치적 시나리오에 대한 생각이 부족합니다. 일단 기술적인 문제를 해결한 뒤에 이런 것들에 대해 고민을 많이 해봐야 합니다.

맥스 테그마크: 얘기만 하는게 아니라 방금 말씀하신 것처럼 실제로 서로 고민하고 협력할 수 있는 자리를 마련하는데 노력하시는 모습이 보기 좋습니다. 실제로 데미스는 이 자리를 마련하는데 도움을 많이 주셨죠. 다음은 닉으로 넘어가볼까요? 당신은 아무런 걱정을 하고 있지 않겠죠? 우리가 집중 해야할 것에 대한 당신의 의견을 얘기해주세요.

닉 보스트롬: 저도 방금 나온 의견과 같은 생각입니다. 기술적인 문제를 해결할 수 있는 전문가를 영입하고 조율하는 것, 조직을 구축하고 신뢰를 쌓는 것. 제도적 관리를 할 수 있는 창의적인 방안을 제공하고 이런 것들을 적용하기 전에 충분히 시도하는 것이 중요하죠.

데이빗 차머스: 저는 개인적으로 의식에 대해 많이 생각합니다. 아실로마 원칙 목록 중 '지각에 대한 주의'에서 '인공지능의 의식에 대한 가정을 피하라'는 원칙에서 충격을 받았습니다. 이 원칙을 보고 저는 이렇게 받아들였습니다. 모든 인간 수준 또는 초 인간 수준 AGI가 의식을 가질 것이라는 가정을 피하라고 말이죠. 개인적으로 저는 이것으로 인해 미래에 실패를 제기할 가능성이 있다고 생각합니다. 인간 또는 초 인간 지능 수준의 AGI를 만들고 그들이 세상에 나와 활동하겠지만 그 누구도 의식을 가지고 있지 않는 세상이 될 가능성이 있다고 생각해요. 많은 사람들은 의미와 가치 있는 삶을 위해 주관적인 인식과 의식이 필요하다는 견해를 가지고 있습니다 그래서 의식이 없는 세상은 긍정적인 결과를 가져올 수 없어요. 끔찍하게 부정적인 결과는 아닐지라도 아무런 결과가 없을 수는 있겠죠. 그래서 저는 그 결과를 피하는 것에 대해 걱정됩니다.

사실 저는 다양한 종류의 인공지능이 의식을 가질 가능성에 대해 상당히 낙관적입니다. 그러나 저희가 이런 가정을 짓는 한 '우리가 AGI를 만들 때, 의식을 가진 존재를 만들려고 하는 것인가?'에 대해 진지하게 생각하는 것이 중요하다고 생각합니다.

맥스 테그마크: 지금까지 열심히 노력해서 연구하고 인공지능을 인간 수준까지 개발했는데 정작 이것을 알아줄 수 있는 의식을 가진 사람이 아무도 없는 상황을 얘기하시는 건가요?

데이빗 차머스: 실존적 리스크는 정말 다양하죠. 그 중 하나는 인류가 멸망하고 인공지능이 세상을 지배하는 것이에요. 몇몇 사람들은 '아 뭐 괜찮아요. 인공지능이 우리의 후손이 되겠죠'라고 얘기하지만 이것보다 더 끔찍한 실존적 리스크는 미래에 이 위대한 사실을 인지할 수 있는 존재가 아무도 없는 것이죠.

맥스 테그마크: 사실 웨인 레그는 지능이 지나치게 발달될 수 있는 가능성에 대해 언급한 적이 있습니다. 그리고 C 언어가 언급되기 까지 긴 시간 동안 강력한 금기도 있었어요. 사실 이 원칙들을 1차적으로 검토할 때 제일 낮은 순위를 차지한 원칙이 의식에 관한 것이었어요. 그래서 우리가..

데이빗 차머스: 설명이 정말 잘못 되었죠.

맥스 테그마크: 용어를 지각으로 변경해서 좀 더 명확하게 설명했어요. 점심때 결과는 더 좋았지만 순위는 여전히 맨 끝이었죠. 하지만 저도 당신의 생각에 어느정도 동의해요.

데이빗 차머스: 지각에 대한 경고는 88%가 동의했어요.

맥스 테그마크: 90%는 아니라서 목록에 올라가지 못했죠. 당신이 사람들의 생각을 바꾸어 더 깊이 생각해볼게끔 만들면 되겠네요. 레이, 하시고 싶은 말씀 있으신가요?

레이 커즈와일: 제가 하려고 한 말은 아니지만 방금 질문에 추가하자면.. AGI를 창조할 때, 사람들은 그저 기계로만 생각하고 의식이 없을 것이라고 단정 짓습니다. 하지만 이것은 슬픈 현실입니다. 우린 올바르게 않는 가정을 하고 있기 때문에 인공지능의 의식과 주관적인 인식을 경계하지 않고 있어요. 제가 하고 싶었던 말은, 인류가 공통적으로 생각하고 있는 3가지 혁명이 있습니다. GNR, 유전학, 생체 기술, 나노 기술 그리고 로봇, 인공지능이기도 하죠. 이에 대한 계획이 항상 있었어요. 수십 년 전 생체 기술에 관한 아실로마 회의가 있었죠. 효과적이었어요. 그리고 나노 기술에 대해서도 비슷한 계획이 있었어요. 하지만 인공지능에 경우 명확하게 입증된 기술적 해결책이 없다는 것이 차이점입니다.

예를 들어 나노 기술같은 경우 기술적 통제가 가능해요. 원칙 중에 자기 복제를 하면 안된다는 원칙이 있습니다. 자기 복제 기능을 사용하지 않고도 많은 수량으로 확장 할 수 없기 때문에 실제로는 현실적이지 않지만 기술 보호는 가능합니다. 인공지능이 우리보다 높은 지능을 갖게 되고 인류와 세계를 파괴하려 하는데 이보다 더 우수한 인공지능이 없다면, 이것은 최악의 상황입니다. 인간은 지구에서 가장 높은 지능을 가지고 있습니다. 그런 인간들이 다른 짐승에게 하는 행위를 관찰해보면 앞서 언급했던 우려가 커질 거예요. 인간은 강아지 같은 애완동물에게 매우 친절 해요. 하지만 농장을 살펴보면 우리보다 낮은 지능을 가지고 있는 짐승에게는 그렇게 친절하지 않아요. 만약 우리보다 더 똑똑한 새로운 유형의 존재가 있다면 우리가 짐승에게 친절하지 않은 것처럼 우리를 똑같이 대할 것입니다. 이런 우려 때문에 이 회의에서 우리가 하고 있는 일이 매우 중요하다고 생각합니다. 또 말씀드리고 싶은 것은, 생체 기술에 관한 아실로마 원칙을 발표했던 것처럼 이번에 지정한 아실로마 원칙도 발표해야 한다고 생각해요. 인공지능 리더들이 실제로 이런 회의를 거쳐서 원칙을 만들었다는 것을 알리는 거죠. 그러면 사람들이 관심을 표하고 원칙에 대해 더 많은 토론을 하겠죠. 그리고 이 원칙을 계속해서 개정해 나가는 거예요. 아실로마의 나노 기술 원칙도 여러 번 개정되었어요. 그러나 저는 실제로 우리가 나서서 이런 원칙을 제시하고 인공지능을 연구하는 커뮤니티가 더 많은 이야기를 나누었으면 좋겠습니다. 그리고 그들로 하여금 우리 연구를 인도할 수 있게 만드는 겁니다. 나노 기술 같은 경우 효과적이었어요.

맥스 테그마크: 바트?

바트 셸맨: 조금 다른 관점을 제시 해보겠습니다. 제가 높게 가지고 있는 관심사 중 하나는 기계가 극도로 영리해지거나 특정 분야에서 똑똑해질 경우입니다. 이런 시기가 와도 과연 인간은 여전히 우리가 무엇을 제안했는지, 어떤 결정을 했는지에 대해 이해할 수 있을까요. 저는 자동화 추리 분야에서 일하고 있는데 지난 20년 동안 상당한 발전이 있었습니다. 일상적으로 해결할 수 있는 변수가 수백가지에서 수백만 가지로 늘었습니다. 이 쪽 분야의 커뮤니티는 이러한 추론 엔진들로부터 해답을 얻고 있어요. 주로 하드웨어 또는 소프트웨어 검증에 대한 해답이죠. 하지만 우리는 더 이상 이러한 해답을 이해할 수 없습니다. 지난 몇 년 동안 인간은 기계를 통해 인간이 이해할 수 있는 해답을 만들 방법을 찾아냈습니다. 따라서 기계의 지능이 인간을 넘어선다 하더라도 기계가 제공해주는 솔루션을 이해할 수 있고 우리가 해결책을 찾을 수 없을지라도 어쩌면 기계가 근접한 설명을 제공할 수 도 있겠죠. 제가 가지고 있는 작고 긍정적인 희망입니다.

맥스 테그마크: 감사합니다. 스투어트?

스튜어트 러셀: 이메일 외에 저를 밤에 잠을 못 이루게 하는 것은 두 가지가 있습니다. 하나는 오용과 악용하려는 자에 대한 문제입니다. 비유하자면 마치 우리가 핵무기를 만들어내고 지구상의 모든 사람들에게 이메일로 전달하는 것과 같아요. '여기 장난감이 있어요. 하고 싶은 데로 하세요.' 이런 것은 어떻게 대응할 수 있습니까? 제 의견으로는 정말로 좋은 해결책이 없다고 판단됩니다. 우리가 해야 할 일 중 하나는 안전한 인공지능을 위한 설계를 매우 명확하고 단순하게 만드는 것입니다. 그리고 우리가 원하는 것 이외의 다른 행위를 못하게 하는 것입니다. 또 한가지는, 인간 수준 인공지능을 도달했지만 인류가 너무 지나치게 의존할 수 있다는 가능성입니다. 점차 단계적으로 인간을 무능력하고 쇠퇴하게 만드는 겁니다. 그리고 리스크를 감소시키는 방법에 대한 제 의견을 말씀드릴게요. 인류가 성장하면서 이루고자 하는 궁극적인 목표는 미래를 바꿀 수 있는, 자유롭게 그릴 수 있는 능력이겠죠. 아마도 인공지능을 통해 이를 수 있는 목표 중에 하나일 겁니다. 그리고 생각해보면, 이 목표를 달성하고 식량 부족같은 문제들이 해결됐을 때 인공지능은 서서히 잊혀질 가능성도 있어요. 미래에 인간은 '인공지능은 성장할 만큼 성장했으니 인류가 성장할 차례'라고 말할 수도 있죠. 오히려 인공지능은 현재보다 덜 관심 받을 수도 있다고 생각합니다.

맥스 테그마크: 좋습니다. 드디어 일론 차례네요. 제가 알기로는 당신은 인공지능에 대한 우려에 대해 이야기 한적이 없는 걸로 알고있어요. 우려되는 사항이 있나요? 우리가 이미 하고 있어야 할 일 같은 것?

일론 머스크: 긍정적인 미래에 대한 생각을 해봤는데요. 어떤 모습일지, 아니면 부정적인 것을 최소한 한 세상은 어떤 모습일지.. 아까 샘이랑 다른 분들이 말씀하셨던 것처럼 인공지능 과정의 끝은 초지능을 확보하는 것 아니면 인류의 멸망이예요. 이 2가지가 미래의 궁극적인 모습이겠죠. 지능은 계속 발전할 거예요. 이것을 가로 막는 유일한 것은 인류를 정지시키거나 파괴하는 것 외에는 없죠. 그래서 우리가 생각해야 할 것은 디지털 초지능을 겸한 세상은 어떨 것인가? 그리고 또 한가지 제가 중요하게 생각하는 것은 우리는 이미 사이보그 입니다. 이미 우리는 기계의 영향을 받고 있어요. 핸드폰 또는 컴퓨터, 기타 등등.. 이미 슈퍼 인간이죠. 사실 우리는 30년전 미국 대통령 보다 더 큰 힘과 능력을 가지고 있어요.

인터넷이라는 능력을 말하는 거예요. 수천 만명의 사람들과 대화를 나눌 수 있고 세계 어디에 있는 사람과도 대화할 수 있죠. 이건 불과 몇 년 전까지만 해도 존재하지 않았던 놀라운 마법의 능력과 같아요. 모든 인간은 슈퍼 인간입니다. 한계는 대역폭입니다. 제약이 있죠. 특히 출력에 제약이 있습니다. 초당 수백 비트 정도 되겠네요. 매우 느립니다. 테라 바이트 수준으로 통신 할 수 있는 컴퓨터랑 비교해보세요. 엄청난 양의 차이입니다. 출력은 시력으로 인해 훨씬 훌륭하지만 이것 또한 크게 향상될 수 있습니다. 저는 DNI를 통해 대역폭 제약의 해결이 이상적인 미래를 위한 제반 조건이라고 생각합니다. 대뇌 피질에 광대역 인터페이스를 적용하는 거죠. 소수의 사람들이 기술을 독점한다면 그들은 지구의 독재자가 될 것입니다. 소수의 사람들이 지구를 통치하는 거죠. 따라서 대역폭의 문제를 해결하는 것이 매우 중요하다고 봅니다. 그리고 인공지능이 모두에게 제공된다면 아마 오늘보다 더 공평한 미래가 될 수도 있겠네요.

맥스 테그마크: 감사합니다. 개인 질문을 받기 전에 드리는 마지막 질문입니다. 일론이 질문으로 잘 이어주는

발언을 하셨네요. 지금까지 우리가 알고 있었던 긍정적인 부분입니다. 실존적 리스크에 대해 정말 많은 이야기를 나누었습니다. 이론적인 이야기는 이쯤에서 그만하고, 티비를 켜서 넷플릭스로 채널을 바꾸어 봅시다. 어떤 과학적 비전이 보이시나요? 티비에서 보여주는 것은 항상 디스토피아죠. 어떠한 이유 때문인지는 몰라도 긍정적인 비전 보다는 두려움과 위협적인 것들이 더 많은 관심을 끌죠. 예를 들어, MIT에 있는 제 사무실로 어떤 학생이 진로 상담을 하러 들어옵니다. 제가 먼저 물어볼 것은 20년 후에 너는 어떤 모습일 것 같니?라는 질문입니다. 그리고 학생은 이렇게 답하겠죠. '뭘 암에 걸릴 수도 있고요. 버스에 치일 수도 있겠죠.' 완전 끔찍한 진로 계획이지 않나요? 저는 이 학생이 좀 더 열정적이었으면 좋겠습니다. 치밀하게 계획을 세우고 피할 수 있는 것은 피하고.. 긍정적인 부분에 대해서 더 듣고 싶네요. 간결하게 한 분씩 바라고 있는, 흥미롭게 생각하는 인공지능에 대한 비전에 대해서 말씀해주시기 바랍니다.

안 탈린: 인공지능의 재미를 늘리고 고통을 줄이는 것이죠. 일라이저의 'Fun Theory' 에서 그는 온갖 인공지능의 부작용으로 인해 만들어진 이상적이지 않은 세상에 대해 상상을 글로 구상했습니다. 소름 끼칠 정도 였죠. 삶에 대한 의미가 없어질 거라는 주장을 했어요. 인공지능의 부작용을 최소화하거나 완화 시킬 수 있는 마련해야 합니다. 지능이 향상될수록 생각의 궤도를 변경시킬 필요가 있습니다.

맥스 테그마크: 감사합니다. 샘, 어떤 것이 기대되나요?

샘 해리스: 이상하게도, 가장 기대되는 부분이 동시에 제가 제일 두려워하는 부분입니다. 지금 이 자리의 취지가 정말 좋은 이유가 인간의 가치에 대한 해답을 깨닫게 해주기 때문이라고 생각합니다. 그리고 푸에르토리코에서 진행된 마지막 미팅에서 누군가가 우리의 철학을 준수하기 위한 데드라인에 놓여 있다라고 얘기했었죠. 인간이 더 나은 가치와 부합하기 위해서 스스로 틀린 것과 올바른 것이 있다는 사실을 인정해야 합니다. 초인공지능의 탄생은 각종 질병을 치료하고 에너지 문제를 해결할 겁니다. 이건 변함 없을 거예요. 하지만 저는 아까 스텐버그가 냈던 의견과 다른 생각입니다. 저는 인류의 흐름이 중단되거나 아니면 인류가 멍청해져서 비생산적인 삶을 사는 것에 대해 걱정하지 않습니다. 제가 기대하고 있는 것은 올바른 가치를 수렴한 초지능입니다. 인간의 가치를 단순히 보호하는 것이 아니라 개선하기 위해 필요한 부분을 제공해주는 초지능이 기대됩니다.

맥스 테그마크: 감사합니다. 데미스 어떻게 생각하세요?

데미스 하사비스: 당연한 말이지만 이런 이유때문에 제가 이쪽 분야에서 일을 하고 있죠. 잘만 하면 인류 역사상 최고의 사건일 겁니다. 어떤 면에서 볼 때, 우리가 가지고 있는 잠재력을 최고로 발휘할 수 있죠. 이미 말씀 드렸듯이, 과학과 의학의 획기적인 발전을 돕는 도구로 사용할 수도 있습니다. 이것은 분명하다고 생각합니다. 제가 인공지능에 관심을 가진 이유는 아마도 이 방에 있는 여러분들과 동일할 겁니다. 우리가 여기서 찾고자 하는지에 대한 큰 질문의 답을 얻기 위해서죠. 우리의 지각을 이해하고, 의식이란 무엇인지, 우주의 본질은 무엇인지, 우리의 목표는 무엇인지. 이런 궁금증에 대한 해답을 얻기 위해서는 인공지능이 필요하다고 저는 생각합니다. 알파고를 활용해 바둑에 대한 궁금증을 찾은 것처럼 미래의 심오한 궁금증을 찾을 수 있다고 생각합니다. 이 과정을 과학과 물리학을 활용해 찾는 겁니다. 아마도 해답을 얻기 위한 과정은 모든 사람들에게 짜릿한 여행이 될 것이라고 생각해요. 아까 잠깐 나누었던 이야기 중 인간이 가장 똑똑한 지능을 가진 존재로서 비교적 낮은 지능을 가지고 있는 동물을 하대한다는 이야기를 다시 짚어보겠습니다. 같이 생각을 해봅시다. 예를 들어 인도에서 살고 있는 호랑이의 분포 범위는 정말 넓습니다. 그리고 거기서 살고 있는 사람들은 가난합니다. 자원이 부족하다는 얘기로. 제가 생각했을 때, 만일 그들에게 풍부한 자원이 주어진다면 의도적으로 호랑이를 죽이려고 하지 않을 겁니다. 아마 그들이 호랑이를 의도적으로 죽이는 이유는 가축을 기르기 위해 필요한 땅을 확보하려는 것이라고 생각합니다. 한 마리의 호랑이가 필요한 영역은 몇 제곱 킬로미터가 필요하겠죠. 하지만 인간 또한 많은 인원을 수용할 수 있는 영역이 필요한 것이죠. 그래서 이런 풍부한 자원 그리고 식량난 같은 문제들을 해결할 수 있다면 인간과 인간 사이의 갈등을 해소하고 제일 중요한 자원 부족을 해결할 수 있겠죠. 비로소 훨씬 더 나은 미래를 맞이 할 수 있어요.

맥스 테그마크: 닉, 아까 당신이 말씀하신 부분 중에서 긍정적인 부분은 조금 짧았는데 이번에는 좀 더 얘기해 주실 수 있나요? 기대되는 부분이 있나요?

닉 보스트롬: 두가지로 말씀드릴 수 있는데요. 하나는 해결하지 못한 문제의 답을 찾는 것입니다. 질병을 치료하는 등 세상에 존재하는 불행과 고통, 부정적인 요소를 제거하는 것이죠. 정말 큰 잠재력이죠.

맥스 테그마크: 닉, 감사합니다. 데이빗?

데이빗 차머스: 저는 인공지능을 활용해 인간을 더 똑똑하게 만들 수 있다는 가능성에 대해 기대가 큼니다. 어떻게 보면 이기적일 수도 있지만, 제 나이는 올해 50입니다. 뇌가 점차 느려지고 바보가 되어가고 있죠. 하지만 딱히 그렇다고 단정 짓기도 애매한 이유는 우리가 이미 사용하고 있는 증강 지능 기술 덕분이죠. 스마트폰 또는 인터넷은 제게 여러 능력을 부여합니다. 과거에는 없었던 확장된 능력이 생기죠. 그리고 앞으로 인공지능은 인간의 능력을 확장 시킬 겁니다. 10년 후에는 사람들은 답러닝이 탑재된 증강현실 글라스를 착용하고 있을 거예요. 물론 제가 60세 정도 되었을 때 저도 하나 필요하겠군요. 그리고 70살이 되면 뇌와 직접 연결시킨 인공지능이 탄생할 수도 있고요. 아니면 인공지능에 뇌를 직접 업로드 한다던가. 그렇게 되면 현재보다 훨씬 더 똑똑해질 수 있겠네요. 딱히 이기적이라고 단정 지을 수는 없겠네요. 그래도 욕심이 나긴 하지만.. 데미스가 인공지능 과학자에 대해 이야기 했었는데 저는 인공지능 철학자를 강조하고 싶습니다. 철학에 대한 궁금증은 매우 어렵습니다 그리고 해답을 찾기에 인간은 너무 멍청하죠. 하지만 인공지능이나 인공지능 기반 인간 아니면 순수 AGI를 활용해 그 한계점을 넘어 각종 철학에 대한 질문의 답을 얻을 수 있다고 생각합니다.

맥스 테그마크: 좋은 말씀 감사합니다. 레이, 당신은 미래의 긍정적인 비전의 선구자로 많은 논문을 발표했죠. 그 중에 당신이 가장 기대하고 있는 것은 어떤 것입니까?

레이 커즈와일: 1만년 뒤로 시간을 되돌려 봅시다. 가서 원시인에게 유익한 미래가 무엇인지 상상해보라는 질문을 해보죠. '어떤 것을 보고 싶어요?' 그럼 이렇게 말하겠죠. '음, 저는 이 불이 영원히 안 꺼졌으면 좋겠어요 그리고 짐승들이 제 동굴 안으로 못 들어오도록 더 큰 돌이 있었으면 좋겠네요.' '다른 건 없나요?' '그 정도면 될 거 같은데요.' '좋은 웹 사이트, 앱, 검색 엔진 이런 건 필요 없나요?' 이번에는 200만년 전으로 돌아가서 영장류와 이야기하는 상상을 해보겠습니다. 말이 통한다면 말이죠. '앞으로 뇌 신피질 기능을 향상 시키는 기술이 나올 거라는데 정말 기대되지 않아요?' 그럼 이렇게 말하겠죠 '그게 무슨 상관이에요?' 그리고 우리는 이렇게 말합니다. '아니 음악이나 유머, 더 많은 것들을 누릴 수 있잔아요' 그럼 당연히 그들의 대답은 '음악이 뭐예요? 유머는 뭐죠?' 라고 하겠죠. 이런 식으로 그들은 이런 개념을 상상조차 할 수 없었습니다. 이와 같이 제가 유추하는 바는 영장류가 상상조차 할 수 없었던 개념이 현실이 된 것처럼 더 심오한 개념을 맞이할 겁니다. 음악과 유머를 넘어서 신피질로 나아가는 거죠. 그렇게 되면 우리는 과거로 돌아가고 싶지 않을 거예요.

맥스 테그마크: 바트 생각은 어때요?

바트 셀먼: 미래를 사전에 정확하게 예측하는 것은 불가능하다고 판단합니다. 개인적으로 수리, 과학 영역에서 큰 발전이 예상 됩니다. 그리고 더 많은 것을 발견하겠죠.

맥스 테그마크: 좋습니다 스투어트의 생각은 어떤가요?

스튜어트 러셀: 제프리 색이 말했던 것과 같이, 많은 사람들에게 삶이란 위대합니다. 하지만 일부 사람들에게 삶이란 위대하지 않아요. 그리고 인공지능의 가장 큰 기대는 후자의 문제를 해결하는 겁니다. 닉이 얘기한 것처럼 긍정적인 요소를 늘려서 부정적인 요소와 균등하게 이루는 것도 좋은 방법이에요. 하지만 제가 생각했을 때는 부정적인 요소를 제거하는 것이 더 중요하다고 생각합니다. 어떻게 했는지는 모르겠지만, 반대편에 앉아있는 데미스가 제 노트를 훑쳐본 것 같네요. 부정적인 요소를 제거하면 사람들의 제로섬 사고방식도 없어지겠죠. 제프리 색이 언급한 것처럼 한 명이 반드시 손해를 봐야 다른 한 명이 이득을

취하는 사고방식은 없어야 합니다. 아마도 이것이 제가 가장 기대하는 부분인 것 같네요. 두번째는 이메일
덜 받는 것입니다.

맥스 테그마크: 드디어 일론 차례입니다. 당신은 미래의 비전에 대해 명확하게 설명한 적이 없는 걸로
알고있는데 지금은 어때요?

일론 머스크: 생각을 많이 해봤어요. 그 결과 두가지로 간추릴 수 있습니다. 머신 브레인의 대역폭 제약을
제거 그리고 인공지능 민주화 입니다. 이 두가지만 갖추면 미래에 대한 걱정이 없어요. 액튼 경의 유명한
인용구 중 이런 것이 있습니다 '권력의 분배는 자유를 심어주고, 권력의 독점은 독재를 낳는다' 인공지능이
모든 인간에게 주어지는 것 그리고 대역폭 제약을 제거하는 것. 이 2 가지만 있다면 미래에 대한 걱정을 할
필요가 없습니다.

맥스 테그마크: 아주 좋습니다. 당이 조금씩 떨어지고 있는 거 저도 알고 있습니다. 그리고 다음 패널을 또
진행해야 하니 신속하게 개인 질문 받겠습니다. 지목되신 분은 이름을 말씀해주시고 패널 한 분을 지목해서
질문해주세요. 요슈아?

질문자 1: 몬트리올에서 온 요슈아 벤시오입니다. 안한테 질문드립니다. 당신의 프레젠테이션에서 영감을
많이 받았습니다. 저는 사람들의 선호도와 가치를 얻는 것에 대해 질문하고 싶습니다. 지금 우리가 하고 있는
연구를 통해 민주주의적 사회를 만들 수 있을까요? 그리고 악용을 방지 하는 방법은 무엇이 있을까요?

안 탈린: 당연하죠. '민주주의 2.0' 아니면 'U.N 2.0' 등등 이런 것도 존재합니다. 제 발표 내용에서도
말씀드렸지만, 사람들은 더 좋은 세상을 원하고 있습니다. 하지만 이렇게 말로만 하는 사람들을 구별해내기가
쉽지 않아요. 셸링 포인트 같이 좀 더 구체적인 기준을 만들 수 있다면 도움이 많이 되겠죠. 민주주의는 수백
년 전부터 시작됐습니다. 사람들의 선호도와 가치를 파악할 수 있는 방법에 대한 지식은 이미 갖춰져
있습니다.

맥스 테그마크: 니콜라스 베르구르엔?

질문자 2: 감사합니다 맥스. 니콜라스 베르구르엔 입니다. 고지식한 질문입니다. AGI 의 능력에 대해 이미
이야기를 많이 나누었고, 일론의 의견에도 동일하는 편입니다. 이 능력이 평등하게 분배되면 다행이죠.
하지만 그렇지 않을 경우, 1년 혹은 10 년이될 수도 있지만, 이 능력을 관리할 수 있는 방법 그리고 누가
관리할 것인지에 대해 의견을 듣고 싶습니다.

맥스 테그마크: 일론 아니면 데미스 두분 중에 한 분이 대답해주시면 좋겠네요.

데미스 하사비스: 전에 제가 말씀드렸던 의견과 관련이 있네요. 재능이 있고 똑똑한 사람들이 더 많이
필요합니다. NIPS(신경정보시스템처리 학회)처럼 인공지능을 연구하는 단체의 규모가 계속해서 커지고
있어요. 하지만 여전히 부족한 건 사실입니다. 거의 모든 사람이 서로의 존재를 알고 있어요. 물론 이걸
서양에 해당되겠죠 그리고 중국 또는 러시아 같은 동양 지역의 상황을 파악하기 어려운 건 당연한 얘기고요.
활동한 정보를 많이 남겨야 합니다.

맥스 테그마크: 감사합니다. 시간이 많이 늦어지고 있는 관계로 마지막 질문 받겠습니다. 이 후에 커피
타임이 마련되어 있으니 추가 질문은 그때 하시면 될 거 같습니다. 마지막 질문은 에릭이 하시죠.

질문자 3: MIT 의 에릭 브릭올프슨 입니다. 아까 일론이 민주주의를 확립하는 것에 대해 말씀하셨습니다. 과연
인공지능은 승자독식효과 즉, 선두자가 독점하려는 경향이 있을까요? 아니면 모두에게 권한이 주어질까요?
그리고 권한을 분배할 수 있는 메커니즘은 어떤 것이 있을까요?

맥스 테그마크: 일론이 대답 해주시겠어요?

일론 머스크: 사회에 위협적인 요소가 발생하면, 정부 기관이라고 말하고 싶진 않지만, 규제가 꼭 필요합니다. 저 또한 규제를 선호하지 않습니다 왜냐하면 (정부의 규제는) 주로 흥을 깨버리는 경향이 있어서요. 하지만 자동차 산업, 항공기 산업에도 규제는 항상 존재합니다. 이외에 의약, 식품도 각종 규제가 존재하죠. 사회적으로 위협을 받을 수 있는 것들은 전부 규제가 있습니다. 인공지능도 예외는 아니죠. 문제는 인공지능이 발달하는 속도에 맞춰 정부가 대응할 수 있는 없는지가 관건입니다. 정부는 확실히 느려요. 그리고 사전에 대책을 마련하지 않는 경향이 있습니다. 다른 산업을 보면, FAA(Federal Aviation Administration) 가 없어지는 것을 원하는 사람은 없을 겁니다. FAA 가 없어진다면 항공기에 대한 규제가 없어지면서 질서도 없어지겠죠. 규제가 존재하는 데는 다 이유가 있습니다. 데미스의 의견에서 덧붙이자면 인공지능 연구 인재를 집중되어있는 곳에서 인공지능이 압도적으로 개발 될 가능성이 높아요. 그리고 세계에 몇 곳 없습니다. 보스톤, 실리콘 벨리 또는 런던이겠죠. 다른 곳을 더 찾아내는 것도 좋지만 실제로 규제 당국이 합리적으로 접근 할 수 있는 곳은 몇 지역이 안됩니다. 그리고 다시 강조하지만, 저는 규제를 좋아하는 편이 아니에요. 골칫거리 입니다. 하지만 때때로 필요한 상황이 있습니다.

맥스 테그마크: 알겠습니다. 이로써 토론을 마치겠습니다. 패널 여러분들께 감사드립니다.