内容概要：Many artificial intelligence researchers expect AI to outsmart humans at all tasks and jobs within decades, enabling a future where we're restricted only by the laws of physics, not the limits of our intelligence. MIT physicist and AI researcher Max Tegmark separates the real opportunities and threats from the myths, describing the concrete steps we should take today to ensure that AI ends up being the best -- rather than worst -- thing to ever happen to humanity.

很多人工智能的研究者预期，在未来十年内，人工智能会在大部分工作上胜过我们。在这样的未来里，我们仅仅受限于物理的定律，而非我们拥有的智慧。麻省理工学院的物理学家和人工智能研究者麦克斯·泰格马克，为我们辨析了未来可能拥有的真正机会和错误认知中的威胁，他为我们说明了今天的我们应当采取哪些步骤，来确保人工智能最终会为人性带来最好的成果——而非糟糕的结局。

www.XiYuSoft.com | **锡育软件**

After 13.8 billion years of **cosmic** history, our universe has **woken** up and become **aware of itself**.	在 138 亿年的历史之后, 我们的宇宙终于觉醒了, 并开始有了自我意识。[00:12]
From a small blue planet, tiny, **conscious** parts of our universe have begun **gazing** out into the **cosmos** with telescopes, discovering something **humbling**.	从一颗蓝色的小星球, 宇宙中那些有了微小意识的部分, 开始用它们的望远镜, 窥视整个宇宙, 从而有了谦卑的发现。[00:21]
We've discovered that our universe is **vastly** grander than our ancestors imagined and that life seems to be an almost imperceptibly small **perturbation** on an otherwise dead universe.	宇宙比我们祖先所想象的 要大得多, 使得生命显得如同渺小的扰动, 小到足以被忽视, 但若没有它们的存在, 宇宙也没了生命。[00:31]
But we've also discovered something **inspiring**, which is that the technology we're developing has the potential to help life **flourish** like never before, not just for centuries but for billions of years, and not just on earth but throughout much of this amazing cosmos.	不过我们也发现了 一些振奋人心的事, 那就是我们所开发的技术, 有着前所未有的潜能 去促使生命变得更加繁盛, 不仅仅只有几个世纪, 而是持续了数十亿年; 也不仅仅是地球上, 甚至是在整个浩瀚的宇宙中。[00:42]

I think of the earliest life as "Life 1.0"	我把最早的生命 称之为 "生命 1.0", [00:59]
because it was really dumb, like **bacteria**, unable to learn anything during its lifetime.	因为它那会儿还略显蠢笨, 就像细菌,在它们的一生中, 也不会学到什么东西。[01:02]
I think of us humans as "Life 2.0" because we can learn, which we in **nerdy**, **geek** speak, might think of as **installing** new software into our brains, like languages and job skills.	我把我们人类称为 "生命 2.0", 因为我们能够学习, 用技术宅男的话来说, 就像是在我们脑袋里 装了一个新的软件, 比如语言及工作技能。[01:08]
'"Life 3.0," which can design not only its software but also its hardware of course doesn't exist yet.	而 "生命 3.0" 不仅能开始设计 它的软件,甚至还可以创造其硬件。 当然,它目前还不存在。[01:19]
But perhaps our technology has already made us "Life 2.1,"	但是也许我们的科技 已经让我们走进了 "生命 2.1",[01:25]
with our artificial knees, **pacemakers** and **cochlear implants**.	因为现在我们有了人工膝盖, 心脏起搏器以及耳蜗植入技术。[01:29]
So let's take a closer look at our relationship with technology, OK?	我们一起来聊聊 人类和科技的关系吧! [01:33]
As an example, the **Apollo** 11 moon mission was both successful and inspiring, showing that when we humans use technology **wisely**, we can accomplish things that our ancestors could only dream of.	举个例子, 阿波罗 11 号月球任务 很成功,令人备受鼓舞, 展示出了我们人类 对于使用科技的智慧, 我们实现了很多 祖先们只能想象的事情。[01:38]

But there's an even more inspiring journey **propelled** by something more powerful than rocket engines, where the passengers aren't just three astronauts but all of **humanity**.	但还有一段更加 鼓舞人心的旅程, 由比火箭引擎更加强大的 东西所推动着, 乘客也不仅仅只是三个宇航员, 而是我们全人类。[01:52]
Let's talk about our collective journey into the future with artificial intelligence.	让我们来聊聊与人工智能 一起走向未来的 这段旅程。[02:03]
My friend Jaan Tallinn likes to **point out** that just as with	我的朋友扬·塔林 （Jaan Tallinn） 常说, 这就像是

English	中文
rocketry, it's not enough to make our technology powerful.	火箭学一样, 只让我们的科技 拥有强大的力量是不够的。[02:08]
We also have to figure out, if we're going to be really **ambitious**, how to steer it and where we want to go with it.	如果我们有足够的 雄心壮志,就应当想出 如何控制它们的方法, 希望它朝着怎样的方向前进。[02:17]
So let's talk about all three for artificial intelligence: the power, the **steering** and the destination.	那么对于人工智能, 我们先来谈谈这三点: 力量,操控和目的地。[02:24]
Let's start with the power.	我们先来说力量。[02:31]
I define intelligence very **inclusively** -- simply as our ability to accomplish complex goals, because I want to include both biological and artificial intelligence.	我对于人工智能的定义非常全面—— 就是我们能够完成复杂目标的能力, 因为我想把生物学 和人工智能都包含进去。[02:33]

English	中文
And I want to avoid the silly carbon-chauvinism idea that you can only be smart if you're made of meat.	我还想要避免愚蠢的 碳沙文主义的观点, 即你认为如果你很聪明, 你就一定有着肉身。[02:44]
It's really amazing how the power of AI has grown recently.	人工智能的力量 在近期的发展十分惊人。[02:52]
Just think about it.	试想一下。[02:56]
Not long ago, robots couldn't walk.	甚至在不久以前, 机器人还不能走路呢。[02:58]
Now, they can do **backflips**.	现在,它们居然开始后空翻了。[03:02]
Not long ago, we didn't have **self-driving** cars.	不久以前, 我们还没有全自动驾驶汽车。[03:05]
Now, we have self-flying **rockets**.	现在,我们都有 自动飞行的火箭了。[03:10]
Not long ago, AI couldn't do face recognition.	不久以前, 人工智能甚至不能完成脸部识别。[03:15]
Now, AI can generate fake faces and **simulate** your face saying stuff that you never said.	现在,人工智能都开始 生成仿真面貌了, 并模拟你的脸部表情, 说出你从未说过的话。[03:19]
Not long ago, AI couldn't beat us at the game of Go.	不久以前, 人工智能还不能在围棋中战胜人类,[03:28]
Then, Google DeepMind's AlphaZero AI took 3,000 years of human Go games and Go wisdom, ignored it all and became the world's best player by just playing against itself.	然后,谷歌的DeepMind推出的 AlphaZero 就掌握了人类三千多年的 围棋比赛和智慧, 通过自己对战的方式轻松秒杀我们, 成了全球最厉害的围棋手。[03:32]
And the most **impressive feat** here wasn't that it crushed human **gamers**, but that it crushed human AI researchers who had spent decades handcrafting **game-playing** software.	这里最让人印象深刻的部分, 不是它击垮了人类棋手, 而是它击垮了人类人工智能的研究者, 这些研究者花了数十年 手工打造了下棋软件。[03:43]

English	中文
And AlphaZero crushed human AI researchers not just in Go but even at chess, which we have been working on since 1950.	此外,AlphaZero也在国际象棋比赛中 轻松战胜了人类的人工智能研究者们, 我们从 1950 年 就开始致力于国际象棋研究。[03:54]
So all this amazing recent progress in AI really **begs** the question: How far will it go?	所以近来,这些惊人的 人工智能进步,让大家不禁想问: 它到底能达到怎样的程度? [04:01]
I like to think about this question **in terms of** this abstract landscape of tasks, where the **elevation** represents how hard it is for AI to do each task at human level, and the sea level represents what AI can do today.	我在思考这个问题时, 想从工作任务中的抽象地景 来切入, 图中的海拔高度表示 人工智能要把每一项工作 做到人类的水平的难度, 海平面表示现今的人工智能所达到的水平。[04:09]
The sea level is rising as AI improves, so there's a kind of global warming going on here in the task landscape.	随着人工智能的进步, 海平面会上升, 所以在这工作任务地景上, 有着类似全球变暖的后果。[04:22]
And the obvious **takeaway** is to avoid careers at the **waterfront** --	很显然,我们要避免 从事那些近海区的工作—— [04:29]
(Laughter) which will soon be **automated** and **disrupted**.	（笑声） 这些工作不会一直由人来完成, 迟早要被自动化取代。[04:33]
But there's a much bigger question as well.	然而同时,还存在一个很大的问题,[04:37]
How high will the water end up rising?	水平面最后会升到多高? [04:40]

English	中文
Will it eventually rise to flood everything, matching human intelligence at all tasks.	它最后是否会升高到淹没一切, 人工智能会不会 最终能胜任所有的工作? [04:43]
This is the definition of artificial general intelligence -- AGI, which has been the holy **grail** of AI research since its	这就成了通用人工智能 （Artificial general intelligence） —— 缩写是 AGI, 从一开始它就是

inception.	人工智能研究最终的圣杯。[04:50]
By this definition, people who say, "Ah, there will always be jobs that humans can do better than machines," are simply saying that we'll never get AGI.	根据这个定义,有人说,"总是有些工作,人类可以做得比机器好。" [04:58] 意思就是,我们永远不会有 AGI。[05:04]
Sure, we might still choose to have some human jobs or to give humans income and purpose with our jobs, but AGI will **in any case** transform life as we know it with humans no longer being the most intelligent.	当然,我们可以仍然 保留一些人类的工作,或者说,通过我们的工作 带给人类收入和生活目标,但是不论如何,AGI 都会转变我们对生命的认知,人类或许不再是最有智慧的了。[05:07]
Now, if the water level does reach AGI, then further AI progress will be driven mainly not by humans but by AI, which means that there's a possibility that further AI progress could be way faster	如果海平面真的 上升到 AGI 的高度,那么进一步的人工智能进展 将会由人工智能来引领,而非人类,那就意味着有可能,进一步提升人工智能水平 将会进行得非常迅速,[05:20]
than the typical human research and development **timescale** of years, raising the **controversial** possibility of an intelligence explosion where **recursively** self-improving AI rapidly leaves human intelligence far behind, creating what's known as superintelligence.	甚至超越用年份来计算时间的 典型人类研究和发展,提高到一种极具争议性的可能性,那就是智能爆炸,即能够不断做自我改进的人工智能 很快就会遥遥领先人类,创造出所谓的超级人工智能。[05:34]

Alright, reality check: Are we going to get AGI any time soon?	好了,回归现实: 我们很快就会有 AGI 吗?[05:51]
Some famous AI researchers, like Rodney Brooks, think it won't happen for hundreds of years.	一些著名的 AI 研究者,如罗德尼·布鲁克斯 (Rodney Brooks),认为一百年内是没有可能的。[05:58]
But others, like Google DeepMind founder Demis Hassabis, are more optimistic and are working to try to make it happen much sooner.	但是其他人,如谷歌DeepMind公司的 创始人德米斯·哈萨比斯(Demis Hassabis) 就比较乐观,且努力想要它尽早实现。[06:03]

And recent **surveys** have shown that most AI researchers actually share Demis's **optimism**, expecting that we will get AGI within decades, so within the lifetime of many of us, which begs the question -- and then what?	近期的调查显示,大部分的人工智能研究者 其实都和德米斯一样持乐观态度,预期我们十年内就会有 AGI,所以我们中许多人 在有生之年就能看到, 这就让人不禁想问—— 那么接下来呢?[06:11]
What do we want the role of humans to be if machines can do everything better and cheaper than us?	如果什么事情机器 都能做得比人好,成本也更低,那么人类 又该扮演怎样的角色?[06:26]
The way I see it, we face a choice.	依我所见,我们面临一个选择。[06:34]
One option is to be **complacent**.	选择之一是要自我满足。[06:37]
We can say, "Oh, let's just build machines that can do everything we can do and not worry about the **consequences**.	我们可以说,"我们来建造机器,让它来帮助我们做一切事情,不要担心后果,[06:39]
Come on, if we build technology that makes all humans **obsolete**, what could possibly go wrong?"	拜托,如果我们能打造出 让全人类都被淘汰的机器,还有什么会出错?" [06:45]

(Laughter) But I think that would be **embarrassingly lame**.	(笑声) 但我觉得那样真是差劲到悲哀。[06:50]
I think we should be more ambitious -- in the spirit of TED.	我们认为我们应该更有野心—— 带着 TED 的精神。[06:55]
Let's **envision** a truly inspiring **high-tech** future and try to steer towards it.	让我们来想象一个 真正鼓舞人心的高科技未来, 并试着朝着它前进。[06:59]
This brings us to the second part of our rocket **metaphor**: the steering.	这就把我们带到了火箭比喻的 第二部分:操控。[07:05]
We're making AI more powerful, but how can we steer towards a future where AI helps humanity flourish **rather than flounder**?	我们让人工智能的力量更强大,但是我们要如何朝着 人工智能帮助人类未来更加繁盛,而非变得挣扎的目标不断前进呢?[07:09]
To help with this, I cofounded the Future of Life Institute.	为了协助实现它,我联合创办了 "未来生命研究所"(Future of Life Institute)。[07:18]
It's a small **nonprofit promoting beneficial** technology use, and our goal is simply for the future of life to exist and to be as inspiring as possible.	它是个小型的非营利机构,旨在促进有益的科技使用,我们的目标很简单,就是希望生命的未来能够存在,且越是鼓舞人心越好。[07:21]
You know, I love technology.	你们知道的,我爱科技。[07:29]

的；可享利益的

English	中文
Technology is why today is better than the Stone Age.	现今之所以比石器时代更好, 就是因为科技。[07:32]
And I'm optimistic that we can create a really inspiring high-tech future ...	我很乐观的认为我们能创造出 一个非常鼓舞人心的高科技未来......[07:36]
if -- and this is a big if -- if we win the wisdom race -- the race between the growing power of our technology and the growing wisdom with which we manage it.	如果——这个 "如果" 很重要—— 如果我们能赢得这场 关于智慧的赛跑—— 这场赛跑的两位竞争者 便是我们不断成长的科技力量 以及我们用来管理科技的 不断成长的智慧。[07:41]
But this is going to require a change of strategy because our old strategy has been learning from mistakes.	但这也需要策略上的改变。 因为我们以往的策略往往都是从错误中学习的。[07:51]
We invented fire, screwed up **a bunch of** times -- invented the fire **extinguisher**.	我们发明了火, 因为搞砸了很多次—— 我们发明出了灭火器。[07:57]
(Laughter) We invented the car, screwed up a bunch of times -- invented the traffic light, the seat belt and the **airbag**, but with more powerful technology like nuclear weapons and AGI, learning from mistakes is a lousy strategy, don't you think?	（笑声） 我们发明了汽车, 又一不小心搞砸了很多次—— 发明了红绿灯,安全带 和安全气囊, 但对于更强大的科技, 像是核武器和 AGI, 要去从错误中学习, 似乎是个比较糟糕的策略, 你们怎么看? [08:02]
(Laughter) It's much better to be **proactive** rather than reactive; plan ahead and get things right the first time because that might be the only time we'll get.	（笑声） 事前的准备比事后的 补救要好得多; 提早做计划,争取一次成功, 因为有时我们或许 没有第二次机会。[08:17]

English	中文
But it is funny because sometimes people tell me, "Max, shhh, don't talk like that.	但有趣的是, 有时候有人告诉我。 "麦克斯,嘘——别那样说话。[08:25]
That's Luddite **scaremongering**."	那是勒德分子（注:持有反机械化, 反自动化观点的人） 在制造恐慌。"[08:30]
But it's not scaremongering.	但这并不是制造恐慌。[08:33]
It's what we at MIT call safety engineering.	在麻省理工学院, 我们称之为安全工程。[08:35]
Think about it: before NASA launched the Apollo 11 mission, they **systematically** thought through everything that could go wrong when you put people on top of explosive fuel tanks and launch them somewhere where no one could help them.	想想看: 在美国航天局（NASA） 部署阿波罗 11 号任务之前, 他们全面地设想过 所有可能出错的状况, 毕竟是要把人类放进 易燃易爆的太空舱里, 再将他们发射上 一个无人能助的境遇。[08:39]
And there was a lot that could go wrong.	可能出错的情况非常多,[08:50]
Was that scaremongering?	那是在制造恐慌吗? [08:52]
That's was precisely the safety engineering that **ensured** the success of the mission, and that is precisely the strategy I think we should take with AGI.	那正是在做安全工程的工作, 以确保任务顺利进行, 这正是我认为处理 AGI 时 应该采取的策略。[08:56]
Think through what can go wrong to make sure it goes right.	想清楚什么可能出错, 然后避免它的发生。[09:04]
So in this spirit, we've organized **conferences**, bringing together leading AI researchers and other thinkers to discuss how to grow this wisdom we need to keep AI beneficial.	基于这样的精神, 我们组织了几场大会, 邀请了世界顶尖的人工智能研究者 和其他有想法的专业人士, 来探讨如何发展这样的智慧, 从而确保人工智能对人类有益。[09:08]

English	中文
Our last conference was in Asilomar, California last year and produced this list of 23 principles which have since been signed by over 1,000 AI researchers and key industry leaders, and I want to tell you about three of these principles.	我们最近的一次大会 去年在加州的阿西洛玛举行, 我们得出了 23 条原则, 自此已经有超过 1000 位人工智能研究者,以及核心企业的 领导人参与签署。 我想要和各位分享 其中的三项原则。[09:17]
One is that we should avoid an arms race and **lethal autonomous** weapons.	第一,我们需要避免军备竞赛, 以及致命的自动化武器出现。[09:31]
The idea here is that any science can be used for new ways of helping people or new ways of **harming** people.	其中的想法是,任何科学都可以 用新的方式来帮助人们, 同样也可以以新的方式 对我们造成伤害。[09:37]
For example, biology and chemistry are much more likely to be used for new medicines or new cures than for new ways of killing people, because **biologists** and **chemists** pushed hard -- and successfully -- for **bans** on biological and **chemical weapons**.	例如,生物和化学更可能被用来 制造新的医药用品, 而非带来新的杀人方法, 因为生物学家和 化学家很努力—— 也很成功地——在推动 禁止生化武器的出现。[09:42]
And in the same spirit, most AI researchers want to **stigmatize** and ban lethal autonomous weapons.	基于同样的精神, 大部分的人工智能研究者也在 试图指责和禁止致命的自动化武器。[09:57]
Another Asilomar AI principle is that we should **mitigate** AI-	另一条阿西洛玛 人工智能会议的原则是, 我们应该

fueled income inequality.	要减轻 由人工智能引起的收入不平等。[10:03]

I think that if we can grow the economic pie **dramatically** with AI and we still can't figure out how to divide this pie so that everyone is **better off**, then shame on us.	我认为,我们能够大幅度利用 人工智能发展出一块经济蛋糕, 但却没能相处如何来分配它 才能让所有人受益, 那可太丢人了。[10:08]
(Applause) Alright, now raise your hand if your computer has ever crashed.	（掌声） 那么问一个问题,如果 你的电脑有死机过的,请举手。[10:18]
(Laughter) Wow, that's a lot of hands.	（笑声） 哇,好多人举手。[10:27]
Well, then you'll appreciate this principle that we should invest much more in AI safety research, because as we put AI **in charge of** even more decisions and infrastructure,	那么你们就会感谢这条准则, 我们应该要投入更多 以确保对人工智能安全性的研究, 因为我们让人工智能在主导 更多决策以及基础设施时,[10:30]
we need to figure out how to transform today's **buggy** and hackable computers into **robust** AI systems that we can really trust, because otherwise, all this **awesome** new technology can **malfunction** and harm us, or get **hacked** and be turned against us.	我们要了解如何将 会出现程序错误以及有漏洞的电脑, 转化为可靠的人工智能, 否则的话, 这些了不起的新技术 就会出现故障,反而伤害到我们, 或被黑入以后转而对抗我们。[10:39]
And this AI safety work has to include work on AI value **alignment**, because the real threat from AGI isn't **malice**, like in silly Hollywood movies, but **competence** --	这项人工智能安全性的工作 必须包含对人工智能价值观的校准, 因为 AGI 会带来的威胁 通常并非出于恶意—— 就像是愚蠢的 好莱坞电影中表现的那样, 而是源于能力——[10:51]

AGI **accomplishing** goals that just aren't **aligned** with ours.	AGI 想完成的目标 与我们的目标背道而驰。[11:03]
For example, when we humans drove the West African black **rhino** extinct, we didn't do it because we were a bunch of evil **rhinoceros haters**, did we?	例如,当我们人类促使了 西非的黑犀牛灭绝时, 并不是因为我们是邪恶 且痛恨犀牛的家伙,对吧? [11:06]
We did it because we were smarter than them and our goals weren't aligned with theirs.	我们能够做到 只是因为我们比它们聪明, 而我们的目标和它们的目标相违背。[11:15]
But AGI is by definition smarter than us, so to make sure that we don't put ourselves in the position of those **rhinos** if we create AGI, we need to figure out how to make machines understand our goals, **adopt** our goals and retain our goals.	但是 AGI 在定义上就比我们聪明, 所以必须确保我们别让 自己落到了黑犀牛的境遇, 如果我们发明 AGI, 首先就要解决如何 让机器明白我们的目标, 选择采用我们的目标, 并一直跟随我们的目标。[11:20]
And whose goals should these be, anyway?	不过,这些目标到底是谁的目标? [11:37]
Which goals should they be?	这些目标到底是什么目标? [11:40]
This brings us to the third part of our rocket metaphor: the destination.	这就引出了火箭比喻的 第三部分:目的地。[11:41]
We're making AI more powerful, trying to figure out how to steer it, but where do we want to go with it?	我们要让人工智能的力量更强大, 试图想办法来操控它, 但我们到底想把它带去何方呢? [11:46]

This is the **elephant in the room** that almost nobody talks about -- not even here at TED -- because we're so **fixated** on **short-term** AI challenges.	这就像是房间里的大象, 显而易见却无人问津—— 甚至在 TED 也没人谈论—— 因为我们都把目光聚焦于短期的人工智能挑战。[11:53]
Look, our species is trying to build AGI, **motivated** by curiosity and economics, but what sort of future society are we hoping for if we succeed?	你们看,我们人类 正在试图建造 AGI, 由我们的好奇心 以及经济需求所带动, 但如果我们能成功, 希望能创造出怎样的未来社会呢? [12:03]
We did an opinion poll on this recently, and I was struck to see that most people actually want us to build superintelligence:	最近对于这一点, 我们做了一次观点投票, 结果很让我惊讶, 大部分的人其实希望 我们能打造出超级人工智能:[12:16]
AI that's vastly smarter than us in all ways.	在各个方面都 比我们聪明的人工智能[12:22]
What there was the greatest agreement on was that we should be ambitious and help life spread into the cosmos, but there was **much less** agreement about who or what should be in charge.	大家甚至一致希望 我们应该更有野心, 并协助生命在宇宙中的拓展, 但对于应该由谁,或者什么来主导, 大家就各持己见了。[12:26]
And I was actually quite **amused** to see that there's some some people who want it to be just machines.	有件事我觉得非常奇妙, 就是我看到有些人居然表示 让机器主导就好了。[12:36]

(Laughter) And there was total disagreement about what the role of humans should be, even **at the most** basic level, so let's take a closer look at possible futures that we might choose to steer toward, alright?	（笑声） 至于人类该扮演怎样的角色, 大家的意见简直就是大相径庭, 即使在最基础的层面上也是, 那么让我们进一步 去看看这些可能的未来, 我们可能去往目的地,怎么样？[12:42]

So don't get be wrong here.	别误会我的意思,[12:55]
I'm not talking about space travel, merely about humanity's **metaphorical** journey into the future.	我不是在谈论太空旅行, 只是打个比方, 人类进入未来的这个旅程。[12:56]
So one option that some of my AI colleagues like is to build superintelligence and keep it under human control, like an **enslaved** god, **disconnected** from the internet	我的一些研究人工智能的同事 很喜欢的一个选择就是 打造人工智能, 并确保它被人类所控制, 就像被奴役起来的神一样, 网络连接被断开,[13:02]
and used to create **unimaginable** technology and wealth for whoever controls it.	为它的操控者创造出无法想象的 科技和财富。[13:13]
But Lord **Acton** warned us that power **corrupts**, and absolute power corrupts absolutely, so you might worry that maybe we humans just aren't smart enough, or wise enough rather, to handle this much power.	但是艾克顿勋爵（Lord Acton） 警告过我们, 权力会带来腐败, 绝对的权力终将带来绝对的腐败, 所以也许你会担心 我们人类就是还不够聪明, 或者不够智慧, 无法妥善处理过多的权力。[13:18]
Also, **aside from** any moral **qualms** you might have about **enslaving** superior minds, you might worry that maybe the superintelligence could **outsmart** us, **break out** and take over.	还有,除了对于奴役带来的优越感, 你可能还会产生道德上的忧虑, 你也许会担心人工智能 能够在智慧上超越我们, 奋起反抗,并取得我们的控制权。[13:31]

But I also have colleagues who are fine with AI taking over and even causing human extinction, as long as we feel the the AIs are our worthy **descendants**, like our children.	但是我也有同事认为, 让人工智能来操控一切也无可厚非, 造成人类灭绝也无妨, 只要我们觉得人工智能 配得上成为我们的后代, 就像我们的孩子。[13:43]
But how would we know that the AIs have **adopted** our best values and aren't just **unconscious zombies tricking** us into anthropomorphizing them?	但是我们如何才能知道 人工智能汲取了我们最好的价值观, 而不是只是一个无情的僵尸, 让我们误以为它们有人性？[13:54]
Also, shouldn't those people who don't want human extinction have a say in the matter, too?	此外,那些绝对不想 看到人类灭绝的人, 对此应该也有话要说吧？[14:04]
Now, if you didn't like either of those two high-tech options, it's important to remember that **low-tech** is suicide from a cosmic perspective, because if we don't **go far** beyond today's technology,	如果这两个高科技的选择 都不是你所希望的, 请记得,从宇宙历史的角度来看, 低级的科技如同自杀, 因为如果我们不能 远远超越今天的科技,[14:10]
the question isn't whether humanity is going to go extinct, merely whether we're going to get taken out by the next killer **asteroid**, supervolcano or some other problem that better technology could have solved.	问题就不再是人类是否会灭绝, 而是让我们灭绝的 会是下一次 巨型流星撞击地球, 还是超级火山爆发, 亦或是一些其他本该可以 由更好的科技来解决的问题。[14:20]

So, how about having our cake and eating it ...	所以,为什么不干脆 坐享其成......[14:30]
with AGI that's not enslaved but treats us well because its values are aligned with ours?	使用非奴役的 AGI, 因为价值观和我们一致, 愿意和我们并肩作战的 AGI？[14:34]
This is the gist of what Eliezer Yudkowsky has called "friendly AI,"	尤多科斯基（Eliezer Yudkowsky） 所谓的 "友善的人工智能" 就是如此,[14:40]
and if we can do this, it could be awesome.	若我们能做到这点,那简直太棒了。[14:44]
It could not only eliminate negative experiences like disease, poverty, crime and other suffering, but it could also give us the freedom to choose from a fantastic new **diversity** of positive experiences -- basically making us the masters of our own destiny.	它或许不会解决负面的影响, 如疾病,贫穷, 犯罪或是其它, 但是它会给予我们自由, 让我们从那些正面的 境遇中去选择—— 让我们成为自己命运的主人。[14:47]
So in summary, our situation with technology is complicated, but the big picture is rather simple.	总的来说, 在科技上,我们的现状很复杂, 但是若从大局来看,又很简单。[15:06]
Most AI researchers expect AGI within decades, and if we just	多数人工智能的研究者认为 AGI 能在未来十年内

bumble into this **unprepared**, it will probably be the biggest mistake in human history -- let's face it.	实现, 如果我们没有事先 准备好去面对它们, 就可能成为人类历史上 最大的一个错误—— 我们要面对现实。 [15:13]
It could enable **brutal**, global **dictatorship** with **unprecedented** inequality, **surveillance** and suffering, and maybe even human extinction.	它可能导致残酷的 全球独裁主义变成现实, 造成前所未有的 不平等监控和苦难, 或许甚至导致人类灭绝。 [15:24]

diversity: n.多样性;差异 **bumble:** vi.犯错误;踉跄;结结巴巴地说话/vt.弄糟;搞坏/n.错误 **unprepared:** adj.无准备的;即席的;尚未准备好的 **brutal:** adj.残忍的; 野蛮的，不讲理的 **dictatorship:** n.专政;独裁权;独裁者职位 **unprecedented:** adj.空前的;无前例的 **surveillance:** n.监督;监视

But if we steer carefully, we could end up in a fantastic future where everybody's better off: the poor are richer, the rich are richer, everybody is healthy and free to **live out** their dreams.	但是如果我们能小心操控, 我们可能会有个美好的未来, 人人都会受益的未来, 穷人变得富有,富人变得更富有, 每个人都是健康的, 能自由地去实现他们的梦想。 [15:32]
Now, **hang on**.	不过先别急。 [15:46]
Do you folks want the future that's **politically** right or left?	你们希望未来的政治 是左派还是右派? [15:48]
Do you want the **pious** society with strict moral rules, or do you an **hedonistic free-for-all**, more like Burning Man 24/7?	你们想要一个有 严格道德准则的社会, 还是一个人人可参与的 享乐主义社会, 更像是个无时无刻 不在运转的火人盛会? [15:52]
Do you want beautiful **beaches**, forests and lakes, or would you prefer to **rearrange** some of those **atoms** with the computers, **enabling** virtual experiences?	你们想要美丽的海滩、森林和湖泊, 还是偏好用电脑 重新排列组成新的原子, 实现真正的虚拟现实? [15:59]
With friendly AI, we could simply build all of these societies and give people the freedom to choose which one they want to live in because we would no longer be limited by our intelligence, merely by the laws of physics.	有了友善的人工智能, 我们就能轻而易举地建立这些社会, 让大家有自由去选择 想要生活在怎样的社会里, 因为我们不会再受到 自身智慧的限制, 唯一的限制只有物理的定律。 [16:07]

live out: 住在外面;活过(某一段时间);实践,实现 **hang on:** 坚持下去;不挂断;握住不放 **politically:** adv.政治上 **pious:** adj.虔诚的;敬神的;可嘉的;尽责的 **hedonistic:** adj.快乐主义者的;快乐论的;快乐主义的 **free-for-all:** n.可自由参加的竞赛;任何人可以参加的竞赛/adj.对任何人开放的 **beaches:** n.[海洋]海滩;[海洋]海滨(beach的复数形式) **rearrange:** vt.重新排列;重新整理 **atoms:** n.[物]原子(atom的复数) **enabling:** adj.授权的/v.使能够;授权给(enable的现在分词)

So the resources and space for this would be **astronomical** -- literally.	所以资源和空间会取之不尽—— 毫不夸张。 [16:18]
So here's our choice.	我们的选择如下:[16:25]
We can either be complacent about our future, taking as an article of blind faith that any new technology is guaranteed to be beneficial, and just repeat that to ourselves as a **mantra** over and **over and over again** as we drift like a **rudderless** ship towards our own **obsolescence**.	我们可以对未来感到自满, 带着盲目的信念, 相信任何科技必定是有益的, 并将这个想法当作 圣歌一般,不断默念, 让我们像漫无目的的船只, 驶向自我消亡的结局。 [16:27]
Or we can be ambitious -- thinking hard about how to steer our technology and where we want to go with it to create the age of **amazement**.	或者,我们可以拥有雄心壮志—— 努力去找到操控我们科技的方法, 以及向往的目的地, 创造出真正令人惊奇的时代。 [16:46]
We're all here to celebrate the age of amazement, and I feel that its essence should **lie in** becoming not **overpowered** but **empowered** by our technology.	我们相聚在这里, 赞颂这令人惊奇的时代, 我觉得, 它的精髓应当是, 让科技赋予我们力量, 而非反过来受控于它。 [16:56]
Thank you.	谢谢大家。 [17:07]
(Applause)	(掌声) [17:09]

astronomical: adj.天文的,天文学的;极大的 **mantra:** n.咒语(尤指四吠陀经典内作为咒文或祷告唱念的);颂歌 **over and over again:** adv.一再地;反复不断地 **rudderless:** adj.无舵的;无指导者的 **obsolescence:** n.[生物]退化;荒废 **amazement:** n.惊异;惊愕 **lie in:** 在于...;睡懒觉;待产 **overpowered:** n.爱情俘虏(歌曲名) **empowered:** adj.获得授权的;允许的/v.允许;授与权力(empower的过去式)