

本期追问

当讲故事的不再是人类，权力的制衡还能靠什么？

AI 比我们更懂自己时，自由意志还剩下多少？

一个没有共享现实的社会，如何维系信任与自我修正？

视频精读 VIDEO STUDY

The next 50 years: humanity, AI, power | Yuval Noah Harari

来源：YouTube 视频 (https://www.youtube.com/watch?v=_V_ed5fuexA)

节目发布：2026-07-30 · 全文 63 段 · 页边注释 63 条 · 生词词组 91 条

目录 CONTENTS

导读 · 讲者与视频总结

第一部分 · 全文对照

01 开场：同时到来的三场革命	0:02
02 认知革命：语言造就人类霸权	2:21
03 民主是对话，技术是地基	5:25
04 AI是能动者，不是工具	8:31
05 法律人格：不可收回的决定	15:07
06 从人类编辑到算法编辑	19:53
07 数据帝国与一键切断开关	22:17
08 AI时代，真相为何节节败退	29:09
09 AI会思考吗？会感受吗？	37:50
10 智慧革命：学会选择愿望	43:59

第二部分 · 核心句型 · 值得模仿的表达

第三部分 · 生词总表 · 复习用

第四部分 · 理解自测 · 8 题

导读 视频总结

本期讲者

尤瓦尔·诺亚·赫拉利 以色列历史学家，耶路撒冷希伯来大学历史系教授，著有《人类简史》《未来简史》《智人之上》，其作品被译成数十种语言、全球销量数千万册。

主持人 本场活动的访谈主持人，具体身份未在文中透露；负责开场引言并围绕AI、权力与真相向赫拉利提问。

第一部分 全文对照 · 附页边注释与生词标注

01 开场：同时到来的三场革命

0:02 主持人

Good evening, everybody. Good evening, Yuval. >> Hey, it's good to be here. >> Looking forward to this conversation very much. I'm just going to give a little brief introduction just before we get stuck into the questions. Cuz I've been fascinated about this topic and looking forward to interviewing you because for most of history, change moved slowly enough that one generation could pass its wisdom to the next with some confidence it would still apply. This is no longer true. In the space of a single lifetime, we are now being asked to absorb three revolutions at once. A new form of intelligence that doesn't just calculate, but also decides. A scramble to power and control that intelligence.

大家晚上好。晚上好，尤瓦尔。>> 嘿，很高兴来到这里。>> 我非常期待这次对话。我先简单做个开场介绍，然后我们再进入正式提问。因为我一直对这个话题非常着迷，也很期待采访你。因为在历史上的大部分时间里，变化的速度足够缓慢，一代人可以把自己的智慧传给下一代，并且有相当的信心，这些智慧依然适用。但现在已经不是这样了。在短短一个人的一生之中，我们如今被要求同时消化三场革命。一种新形式的智能，它不仅会计算，还会做决定。一场对这种智能的权力与控制权的争夺战。

主持人开场提出全场框架：三场同时发生的革命——会决策的新智能、争夺其控制权的权力竞赛、比我们更懂我们的数据爆炸。后文所有话题都围绕这三条线展开。

get stuck into *phr.*

开始认真投入做（某事）；英式口语，常指开始正题或大快朵颐

scramble /'skræmbəl/ *n.*

争夺，抢夺（a scramble for/to power 权力争夺战）

0:49

And an explosion of data that lets that intelligence know us better than we know ourselves. Yuval Noah Harari's work from *Sapiens* to *Homo Deus* to *Nexus* and the numerous talks, lectures, and interviews he's conducted has offered the same argument just in different clothes. Humans don't run on truth, we run on shared stories. And whoever controls the story controls the species. What's different now is for the first time the storyteller may not be human. AI can generate belief, manufacture evidence, and forge intimacy at a scale no emperor, church, or party ever could.

还有数据的大爆炸，让这种智能对我们的了解超过我们对自己的了解。尤瓦尔·诺亚·赫拉利的著作，从《人类简史》到《未来简史》再到《智人之上》，以及他做过的众多演讲、讲座和访谈，其实都在以不同的形式表达同一个论点：人类不是靠真相运转的，我们靠的是共同的故事。而谁掌控了故事，谁就掌控了整个物种。如今不同的是，历史上第一次，讲故事的人可能不再是人类。AI 能够制造信念、伪造证据、营造亲密感，其规模之大是任何皇帝、教会或政党都无法企及的。

1:34

This is the turning point. Not simply that the technology is powerful, but the old ways that held power in check, a free press, a shared reality, institutions built to self-correct, were designed for a world where only humans held the pen. Tonight's conversation asks what it takes to carry those checks forward or what the new ones we will need to invent for a future world that no longer we fully author alone. Well, that's the precedent that we're going to look at today, Yuval. But actually before we get into it being too heavy, wanted to start with a slightly light and fun question.

这就是转折点。问题不仅仅在于这项技术很强大，而在于过去那些制衡权力的方式——自由的新闻界、共同的现实、为自我纠错而建立的制度——都是为一个只有人类执笔的世界设计的。今晚的对话要探讨的是，我们需要做什么才能把这些制衡延续下去，或者我们需要发明哪些新的制衡机制，来应对一个不再由我们独自书写的未来世界。嗯，这正是我们目前面临的先例。我们今天探讨的内容，尤瓦尔。不过其实在进入太沉重的话题之前，我想先从一个稍微轻松有趣的问题开始。

赫拉利三部曲：《人类简史》(Sapiens, 2011)讲人类如何靠虚构故事合作崛起，《未来简史》(Homo Deus, 2016)谈数据主义与人类未来，《智人之上》(Nexus, 2024)专论信息网络与AI。

forge /fɔːrdʒ/ v.

锻造；缔造（关系）；伪造。此处 forge intimacy 指制造亲密感

intimacy /ˈɪntɪməsi/ n.

亲密，亲密感

主持人点出全场核心问题：自由媒体、共同现实、自我纠错制度等权力制衡机制，都是为「只有人类执笔」的世界设计的，AI时代需要延续或重造这些制衡。

held power in check *phr.*

制衡权力 (hold sth. in check 抑制、约束)

precedent /ˈprezɪdənt/ n.

先例，前例（法律与历史语境常用）

02 认知革命：语言造就人类霸权

2:21 尤瓦尔·赫拉利

If you could travel back to any point in history, when would you choose and why based on wanting to understand first hand some of the stories that society could tell you? >> Well, I'd probably want to go back to the Stone Age and to the cognitive revolution around 50, 60, 70,000 years ago. I probably wouldn't be able to survive for more than a few days because I don't know how to gather or escape predators or whatever, but it is the most fascinating moment in human history, the one we least understand. We know that within a very short evolutionary time span, we shifted from being a relatively insignificant animal to being the masters of the planet.

如果你可以穿越回历史上任何一个时刻，你会选择什么时候，又是为什么？前提是你想亲身了解一些社会能告诉你的故事。>> 嗯，我可能会想回到石器时代，回到大约五六万到七万年前的认知革命时期。我可能撑不了几天就活不下去了，因为我不懂怎么采集食物，也不会躲避捕食者之类的，但那是人类历史上最迷人的时刻，也是我们最不了解的时刻。我们知道，在极其短暂的演化时间里，我们从一种相对微不足道的动物变成了这个星球的主宰。

3:11

And we are not sure how we did it. The best speculation or the best theory is that we did it with language. That this is the moment when humans acquire the ability to produce and communicate with sophisticated language and that this was the basis for everything else we did. That all the enormous structures we've built on Earth, churches, states, trade networks, financial systems, they're all ultimately made from words. The words in the law books, in the religious scriptures, in the ledgers.

而我们并不确定这是怎么做到的。最好的推测，或者说最好的理论，是我们靠语言做到的。也就是说，那正是人类获得了一种能力的时刻——能够创造复杂精妙的语言并用它来交流，而这正是我们后来所做一切的基础。我们在地球上建立的所有庞大结构——教会、国家、贸易网络、金融体系——它们归根结底都是由词语构成的。法律书里的词语、宗教经典里的词语、账本里的词语。

认知革命(Cognitive Revolution)是《人类简史》的核心概念：约5-7万年前智人在认知能力上发生突变，获得复杂语言，随后走出非洲并淘汰其他人种。

赫拉利的标志性论点：国家、教会、金融体系等庞大结构本质上由词语构成（法律条文、经文、账本），是「互为主体的现实」——只存在于人类共同相信之中。

speculation /ˌspek.jəˈleɪʃən/ *n.*
推测，猜想（无实证的理论）

sophisticated /səˈfɪstɪkeɪtɪd/
adj.
复杂精妙的，高度发达的

scriptures /ˈskɪptʃəz/ *n.*
宗教经文，经典

ledgers /ˈledʒəz/ *n.*
账本，分类账（会计术语）

3:51

And this is especially important to understand today when something new is mastering language, which is AI, maybe for the first time in history, there will be something on the planet that is better with words than we are. And if this is the operating system of civilization, what happens when something else, something alien, takes control of the operating system of civilization? One last thought is that perhaps AI is not a machine that learns language. Perhaps it is language itself liberating itself from its dependence on these animals, on these apes.

而这一点在今天尤其重要，因为现在有一种新事物正在掌握语言，那就是AI。也许这是历史上第一次，这个星球上将出现一种比我们更擅长运用词语的东西。而如果语言是文明的操作系统，那当另一种东西、一种异类的东西掌控了文明的操作系统时，会发生什么？最后还有一个想法：也许AI并不是一台学习语言的机器。也许它就是语言本身，正在把自己从对这些动物、这些猿类的依赖中解放出来。

「语言是文明的操作系统」是理解全场的关键比喻：AI掌握语言，等于掌握了操作系统本身，因此其影响远超以往任何技术。

alien /'eɪliən/ *adj.*

异类的，陌生的；非人类的（此处形容AI是「异质」存在）

4:43

That yes, we created it. It was our most magnificent and important tool. But now language liberate itself from human beings. And what will happen as we move forward is language developing and spreading maybe throughout the universe without us. >> Yeah, fascinating. I mean, that's looking ahead at the future of language being taken away from us. But can we look back again at some of the technological advances within language? We've gone through writing, the printing press, for example. AI obviously is a massive leap >> Yes.

没错，是我们创造了它。它曾是我们最辉煌、最重要的工具。但现在，语言正把自己从人类手中解放出来。而接下来会发生的，就是语言自行发展、传播，也许遍布整个宇宙——不再需要我们。>> 是啊，太有意思了。我是说，这是在展望未来，语言被从我们手中夺走。但我们能不能再回头看看语言领域的一些技术进步？比如我们经历了文字、印刷术。AI 显然是一次巨大的飞跃>> 是的。

赫拉利提出一个激进的思想实验：AI或许不是「学会语言的机器」，而是语言本身脱离人类载体、独立演化——这一想法在第56段被再次呼应。

03 民主是对话，技术是地基

5:25

>> ahead in that. You've used the term that it could cause earthquakes in democracies >> Mhm. >> in how it's in how it powerful it can be with languages. Why could it cause earthquakes in democracies, say, compared to any other form of government or perhaps totalitarian forms of government? >> Because democracy in essence is a conversation. Whereas dictatorship is a dictate. You know, there is one person dictating everything. Democracy is about lots of people having a conversation to decide what to do. To have peace or war, to have raised taxes or lowered taxes. Now, conversations are based on the communication technology, the information technology available to people at the time.

>> 在这方面。你用过一个说法，说它可能在民主制度中引发地震>> 嗯。
>> 就凭它在语言上能有多强大。为什么它会在民主制度中引发地震？比方说，相比其他任何形式的政府，或者极权主义形式的政府？>> 因为民主在本质上是一场对话。而独裁则是一种命令。你知道，就是一个人独断一切。民主是很多人通过对话来决定该做什么。是和平还是战争，是加税还是减税。那么，对话是建立在当时人们可用的通信技术、信息技术之上的。

6:17

And what we see in history, that we do not know, at least I don't know, of a single case of a large-scale democracy before the modern era. We have plenty of examples of small-scale democracies in the ancient and medieval worlds. We have city-states like ancient Athens or medieval Florence. We have lots of tribes and towns and villages run democratically. It seems that at least in the Stone Age, democracy was the most common system of government. There are a few tribes with a single dictator controlling everybody.

而我们在历史上看到的是——我们不知道，至少我不知道——在现代之前有过任何一个大规模民主的例子。我们有很多古代和中世纪世界里小规模民主的例子。我们有城邦，比如古雅典或中世纪的佛罗伦萨。我们有很多部落、城镇和村庄是民主运作的。看起来至少在石器时代，民主是最常见的政体。也有少数部落是由一个独裁者控制所有人的。

定义式对比：民主的本质是「对话」(conversation)，独裁的本质是「命令」(dictate)。由此推出：对话依赖信息技术，信息技术的变革必然震动民主。

dictate /'dɪktet/ *n.*

命令，指令。注意名词重音在前，动词 /dɪk'tet/ 重音在后

历史论据：现代之前没有大规模民主的实例，只有小规模民主——古雅典城邦、中世纪佛罗伦萨、部落村镇。石器时代民主反而是最常见的政体。

medieval /,medi'i:vəl/ *adj.*

中世纪的

city-states /'sɪti steɪts/ *n.*

城邦（古希腊雅典、中世纪佛罗伦萨等自治城市国家）

6:53

But, as human systems grew larger, and you had kingdoms with millions of people spread over thousands of kilometers, it was impossible to hold a conversation anymore. You could do it in a small city like Athens. Everybody can gather together in the main square and discuss. But, how do you do it if even in a relatively small kingdom like say Portugal in the Middle Ages? Can't be done. And therefore, we don't have any example of a large-scale democracy before the rise of modern information technology. First with print, newspapers, then radio and television and the internet.

但是，随着人类体系越来越大，王国拥有呢数百万人口，分布在数千公里的疆域上，就不可能再进行对话了。在雅典那样的小城市里你可以做到。所有人可以聚集在中心广场上讨论。但是，你要怎么做到呢，哪怕是在一个相对较小的王国，比如中世纪的葡萄牙？根本做不到。因此，在现代信息技术兴起之前，我们没有任何大规模民主的例子。首先是印刷、呢、报纸，然后是广播、电视和互联网。

7:39

And these technologies are not kind of a side dish. Like you have the democratic banquet and yes, we also have these newspapers and in radio. They are the basis. They provide the means for millions of people to hold a meaningful conversation in real time about whether to go to war or make peace. And every time there is a major change in these information technologies, you have an earthquake in the building that is built on top of that, which is democracy. And this is what we have been seeing over the last 10 years for instance, as social media became maybe the most important information technology and causing a massive earthquake in democracies all over the world as a result.

而这些技术并不是什么配菜。不是说你摆了一桌民主的盛宴，然后顺便还有这些报纸和广播。它们是根基。它们提供了让数百万人就是否开战或讲和进行实时、有意义对话的手段。而每当这些信息技术发生重大变革时，你就会看到一场地震，发生在建立于其上的那座建筑里，也就是民主。而这正是我们过去十年来一直看到的，比如，社交媒体呢，成为了也许是最重要的信息技术，呢，结果在全世界的民主国家引发了一场巨大的地震。

规模是关键变量：数百万人分布数千公里时对话无法进行，所以中世纪葡萄牙也做不到民主。印刷、报纸、广播、电视、互联网才使大规模民主成为可能。

「不是配菜而是根基」：信息技术不是民主的附属品，而是其存在条件。过去十年社交媒体成为主导信息技术，因此在全球民主国家引发「地震」。

side dish *n.*

配菜；喻指附属品、次要之物

banquet /'bæŋkwɪt/ *n.*

宴会，盛宴

04 AI是能动者，不是工具

8:31 主持人

>> Yeah, one thing you said there was the choice between going to war or making peace. And actually I want to get into the definition of AI because a lot of people call it a tool, but you argue that it's absolutely an agent. >> Yes. >> And in fact you say it's a knife that can decide by itself whether to cut salad or commit murder. And I have to say it would be a bizarre cooking show if that was the case where knives had that ability to cut salads or commit murder. So it is a it's a dangerous thought. Can we expand on that?

>> 是啊，你刚才提到的一点是在开战与讲和之间做选择。其实我想聊聊AI的定义，因为很多人把它叫做工具，但你认为它绝对是一个智能体（agent）。>> 是的。>> 而且实际上你说它是一把可以自己决定是切沙拉还是去杀人的刀。我得说，那会是一档非常诡异的烹饪节目，如果真是那样，刀子有那种能力自己决定切沙拉还是杀人。所以这是个危险的想法。我们能展开讲讲吗？

9:05 尤瓦尔·赫拉利

Are we choosing the right narrative at the moment with AI by saying it's a tool that humans can still use and that humans are in the >> But that's the biggest difference. >> just need to say it's agent and they can do the >> agent. I mean, if it's not an agent, then it's not an AI and all the hundreds of billions of dollars that are being poured into its development will be wasted. Because the expectations, what people expect this technology to do, the reasons they invest hundreds of billions of dollars is because they think it will be an agent, not a tool.

我们现在对 AI 选择的叙事对吗？说它是一个人类仍然可以使用的工具，人类还处于——>> 但这就是最大的区别。>> 只需要说它是智能体，它们就能做那些——>> 智能体。我是说，如果它不是智能体，那它就不是 AI，而所有投入其研发的数千亿美元呢，都会打水漂。因为那些期望——人们期望这项技术做到的事，他们投入数千亿美元的原因——就是他们认为它会是一个智能体，而不是工具。

「会自己决定切沙拉还是杀人的刀」是赫拉利常用的比喻，用来说明AI与一切既往技术的本质区别：它是能动者(agent)，不是工具(tool)。

bizarre /bɪˈzɑːr/ *adj.*

古怪的，离奇的

commit murder *phr.*

杀人，犯谋杀罪（commit + 罪行名词的固定搭配）

反驳「AI只是工具」的釜底抽薪论证：数千亿美元的投资逻辑本身就建立在AI是自主能动者的预期上——如果只是工具，这些钱就白花了。

poured into *phr.*

（资金等）大量投入（pour money into 砸钱于）

9:37

The you listen to the to the narrative coming out of places like Silicon Valley, and there is an inherent contradiction in the in the narrative. They tell us, "We will create a god, and it will be our slave." And this doesn't make sense. If it is a god, it cannot remain a slave. And if it is a slave, it means it has no god-like abilities. The assumption of most of the leading companies, and most of the leading experts also outside the companies, is that they're correct about the first half of the sentence. They are creating a god.

「我们将创造一个神，而它会是我们的奴隶」——赫拉利指出硅谷叙事的逻辑矛盾：神不可能永远是奴隶，奴隶则谈不上有神力，两者不可兼得。

inherent contradiction *phr.*

内在矛盾 (inherent /ɪnˈhɪərənt/
固有的、与生俱来的)

呃，你听从硅谷这类地方传出来的叙事，那套叙事里有一个内在的矛盾。他们告诉我们：“我们要创造一个神，而它会是我们的奴隶。”这说不通。如果它是神，它就不可能一直是奴隶。而如果它是奴隶，那就意味着它💎💎有神一般的能力。而呃，大多数领先公司的假设，以及公司之外大多数顶尖专家的假设，是他们对这句话的前半部分是正确的。他们正在创造一个神。

10:21

A god in the sense of something with superhuman abilities. Now, and this is in essentially an agent. The difference between tool and agent, a tool is something that cannot make decisions by itself. We have to decide what to do with it, and cannot invent new ideas by itself. The printing press was a tool. Mhm. The printing press by itself, Gutenberg brings print technology to Europe in the 15th century. The printing press cannot decide by itself, mhm, "I want to print the Bible, and tomorrow I'll print the Quran, and next week I'll print Confucius." No, Gutenberg decides, "I'll print the Bible," and the printing press is just a tool to realize the decision of Gutenberg.

工具的定义通过古腾堡展开：约1450年古腾堡将活字印刷引入欧洲，但决定印《圣经》的是古腾堡本人——印刷机既不能自主决策，也不能发明新想法。

一个神，意思是某种拥有超人能力的东西。那么，而且，这本质上就是一个能动者 (agent)。工具和能动者的区别在于，工具是一种不能自己做决定的东西。我们必须决定用它来做什么，它也不能自己发明新想法。印刷机就是一种工具。嗯。印刷机本身，古腾堡在15世纪把印刷技术带到了欧洲。印刷机不能自己决定，嗯，“我想印《圣经》，明天我要印《古兰经》，下周我要印孔子的书。”不，是古腾堡决定“我要印《圣经》”，而印刷机只是一个工具，用来，呃，实现古腾堡的决定。

11:14

Similarly, the printing press cannot invent a new idea. The printing press cannot invent the radio. It cannot invent a book. You need humans to write the book, and then the printing press copies it. Similarly, an atom bomb is a tool. An atom bomb in the 1940s could not decide by itself whether to bomb Hiroshima or to bomb Tokyo. You needed a human to decide, "Let's bomb Hiroshima." And the atom bomb could not invent by itself the hydrogen bomb. Impossible. An agent, in contrast, is something that can make decisions by itself and invent new things by new ideas by itself. AI is an agent. An AI weapon can decide what to bomb.

同样，印刷机不能发明新想法。印刷机不能发明收音机。它不能创作一本书。你需要人类来写书，然后印刷机把它复印出来。同样，原子弹也是一种工具。20世纪40年代的原子弹不能自己决定是轰炸广岛还是轰炸东京。你需要一个人来决定："我们轰炸广岛吧。"而且原子弹不能自己发明出氢弹。不可能。相比之下，能动者是一种能够自己做决定、自己发明新事物、新想法的东西。AI就是一个能动者。AI武器可以决定轰炸什么目标。

12:07

And it can invent the next weapon. Similarly, there is a lot of interest now in creating AI agents in the financial and economic sphere. Creating non-human corporations. Imagine a corporation that manages money and employees and has no humans as executives or shareholders or trustees. Only AIs. This is a new type of agent that never existed before in history. Now, is it good? Is it bad? Leave it aside. The first thing to grasp is that this is unprecedented. We have never in we had other agents around us like horses and cows and chickens and birds, lots of animal agents that can make their decisions.

它还能发明下一代武器。同样，现在有很多人有兴趣，呃，在金融和经济领域创造AI能动者。创造非人类的公司。想象一个公司，它管理资金和员工，而且，呃，没有任何人类担任高管、股东或受托人。只有AI。这是历史上从未存在过的一种新型能动者。那么，它是好是坏？先放到一边。首先要理解的是，这是史无前例的。我们从来没有——我们身边曾经有过其他能动者，比如马、牛、鸡和鸟，很多能自己做决定的动物能动者。

原子弹例子强化论证：1945年轰炸广岛的决定由人类做出，原子弹也不能自己发明氢弹。而AI武器既能自选目标，又能设计下一代武器——两条工具的边界都被突破。

「非人类公司」：完全由AI担任决策者、没有人类高管股东的公司。赫拉利强调先别急着评判好坏，要先意识到这是历史上从未有过的能动者类型。

sphere /sfiə/ *n.*

领域，范围 (the financial sphere 金融领域)

shareholders /'ʃerhəʊldəz/ *n.*

股东

trustees /trʌ'sti:z/ *n.*

受托人，托管人（法律术语）

unprecedented /ʌn'presɪdntɪd/ *adj.*

史无前例的，空前的

12:59

But, we never encountered an agent that understands our language and that is better than us at things like finance or law or religion. Again, we used language over thousands of years to create the control systems of the world. The control systems of the world are these huge bureaucratic networks ultimately based on language like banks, like the stock market, like the Catholic Church, like the legal system. We never really liked these democrat— these bureaucratic systems. Humans generally don't like bureaucracy very much, but we can't live without them.

但是，我们从来没有遇到过一个能理解我们语言的能动者，而且它在某些方面比我们更强，比如金融，或者法律、宗教。再说一遍，我们用了几千年的语言来创造这个世界的控制系统。世界的控制系统是这些庞大的官僚网络，最终建立在语言之上，比如银行、比如股市、比如天主教会、比如法律体系。我们从来都不太喜欢这些民主——这些官僚系统。人类一般都不太喜欢官僚体制，但我们又离不开它们。

13:44

They're the basis for our power. AI is a bureaucratic native. What is likely to happen, what is already be happening, don't expect to see killer robots running in the streets shooting people. This is not how AI will take over the world. It will be the AI bureaucrats. We control the world because of these networks of bureaucracy, and we are not very good at it. We only control them because there was nobody else on the planet that could take them from us. We are bad at finance, but the horses are even worse.

它们是我们权力的基础。AI是官僚体制的原住民。可能会发生的，其实已经在发生的是——别指望看到杀手机器人在街上跑来跑去朝人开枪。AI不会以这种方式接管世界。接管世界的会是AI官僚。我们之所以控制着世界，是因为这些官僚网络，而我们其实并不擅长运作它们。我们之所以能控制它们，只是因为地球上没有别人能从我们手里夺走它们。我们不擅长金融，但马更不擅长。

银行、股市、天主教会、法律体系被统称为「世界的控制系统」——庞大的、基于语言的官僚网络。人类讨厌官僚体制，却靠它维持对世界的支配。

bureaucratic /ˌbjʊərəˈkrætɪk/
adj.
官僚的，官僚体制的

全场金句之一：「AI是官僚体制的原住民」。AI接管世界的方式不是街头杀手机器人，而是接管官僚系统——人类掌权只是因为此前地球上没有更强的竞争者。

bureaucratic native *phr.*
官僚体制的「原住民」；仿 digital native（数字原住民）的造词

take over *phr. v.*
接管，接手控制

14:21 主持人

We are bad in managing the legal system, but the chickens are even worse. Now, if somebody comes along that is better than us at bureaucracy, at finance, at law, that's the question. What will happen to us at that moment? >> Well, let's look at the legal side of it a bit and the fact that AI as an agent currently doesn't have any legal personhood. We don't give it the same human legalities, do we? But, you've said before that if we don't decide where AI sits from a legal standpoint when we're talking about autonomous vehicles, etc., someone else will make that decision for us in 10 years and we're out of that power loop. That's a pretty scary thing.

我们不擅长管理法律体系，但鸡更不擅长。那么，如果出现了一个比我们更擅长官僚运作、金融、法律的存在，问题就来了。到那一刻我们会怎么样？>> 好，我们稍微谈谈这里面的法律层面，以及这样一个事实：AI作为能动者，目前不具备任何法律人格。我们并没有赋予它和人类同样的法律地位，对吧？但你以前说过，如果我们不决定AI在法律上的定位——比如在谈论自动驾驶汽车等等的时候——那么10年后就会有别人替我们做这个决定，而我们就被排除在权力圈之外了。这挺可怕的。

「我们不擅长金融，但马更差；不擅长法律，但鸡更差」——幽默背后是冷峻的问题：当出现比人类更擅长官僚、金融、法律的存在，人类地位如何？

legal personhood *n.*

法律人格（法律上作为权利义务主体的资格）

autonomous /ɔ:'tɑ:nəməs/ *adj.*

自主的，自治的（autonomous vehicles 自动驾驶汽车）

05 法律人格：不可收回的决定

15:07 尤瓦尔·赫拉利

But, I want to pull back a bit here cuz the one decision on the table right now that you believe once made humanity will not be able to take back? Is it that legal personhood of AI or is there something of AI, or is that something >> Yes, I would point at that. That there is a lot of attention on the technological race to develop the technical side of AI, to make it faster, more efficient, more powerful. But I would focus also on the legal and political aspect, and especially on the question whether we grant legal personhood to AIs. Now, what is legal personhood?

不过，我想稍微回到一个问题，因为现在摆在桌面上有一个决定，你认为一旦做出，人类将无法收回？是不是就是AI的法律人格问题，还是另有其他的东西，还是别的什么？>> 是的，我会指向这个问题。现在有很多注意力都放在技术竞赛上，发展AI的技术层面，让它更快、更高效、更强大。但是我还想聚焦在法律和政治层面，尤其是这个问题：我们是否要赋予AI法律人格。那么，什么是法律人格？

赫拉利认为最不可逆的决定不是技术层面，而是法律政治层面：是否赋予AI法律人格。舆论聚焦技术竞赛时，真正的分水岭可能在立法机构悄然通过。

15:52

A person in the legal system is someone who, for instance, can open a bank account. Someone who can be a player, an independent player, in the financial system, make investments. Someone who can be a player in the legal system, for instance, sue you in court. Now, until today, there are just two kinds of legal persons in the world. You had human beings, of course, we are persons. I can open a bank account, I can sue you in court, I can donate money to a politician. I'm a legal person, but most legal systems in the world also recognized in the modern era another kind of legal person, which are corporations.

法律体系中的“人”，是指比如说，可以开设银行账户的主体。一个可以在金融体系中作为独立参与者、进行投资的主体。一个可以在法律体系中作为参与者的主体，比如说，在法庭上起诉你。那么，直到今天，世界上只有两种法律人。首先当然是人类，我们是“人”。我可以开银行账户，可以在法庭上起诉你，可以给政治家捐款。我是一个法律人，但世界上大多数法律体系在现代还承认另一种法律人，那就是公司。

16:38

According to not all, but most legal systems, corporations, like Google or Facebook or Toyota or Mercedes-Benz, they are legal persons. The corporation can open a bank account. The corporation can sue you in court. The corporation can make a political donation. Now, until today, this was legal fiction. Because all the decisions of the corporation were actually made by some human being. If Google decides to acquire another company, who actually made the decision? Not Google, but the human executives, the human shareholders, the human engineers made a decision. If Google decides to sue you in court, it was some human employed by Google.

按照大多数（虽然不是全部）法律体系，公司——比如谷歌、Facebook、丰田或梅赛德斯-奔驰——它们都是法律人。公司可以开设银行账户。公司可以在法庭上起诉你。公司可以进行政治捐款。不过，直到今天，这都是一种法律拟制。因为公司的所有决定实际上都是由某个人类做出的。如果谷歌决定收购另一家公司，实际上是谁做的决定？不是谷歌，而是人类高管、人类股东、人类工程师做的决定。如果谷歌决定在法庭上起诉你，那也是谷歌雇佣的某个人做的。

法律人(legal person)的操作性定义：能开银行账户、独立投资、在法庭起诉、政治捐款。目前世界上只有两类法律人：自然人和公司。

sue /su:/ v.

起诉，控告 (sue sb. in court)

「法律拟制」(legal fiction)：公司虽有法律人格，但一切决定实际由人类高管、股东做出——

「从来不存在一个真正的谷歌」。这是AI人格与公司人格的根本区别。

legal fiction n.

法律拟制：法律上「假定为真」的虚构安排，如公司人格

17:26

There was no real Google. Now there can be. If we grant AIs legal personhood, and just a month ago Argentina, the government of Argentina announced that it is going to grant legal personhood to AIs so they can start having operating non-human corporations. Corporations, again, that can employ people, can sue you in court, can do anything a corporation does, and has no humans. Now, 10 years ago, this would just be a crazy idea because if there are no humans, who makes the decisions? Yeah.

从来不存在一个真正的“谷歌”。而现在可以有了。如果我们赋予AI法律人格——就在一个月前，阿根廷，阿根廷政府宣布将要赋予AI法律人格，这样它们就可以开始运营非人类的公司。这些公司，同样，可以雇人，可以在法庭上起诉你，可以做公司能做的一切，而且没有任何人类。那么，10年前，这只会是一个疯狂的想法，因为如果没有人类，谁来做决定呢？对吧。

18:04

You can pass a law in parliament that you can have a non-human corporation, but it can't function. Now it's technically possible. AI can manage, for instance, a bank account, and decide to invest here, and buy this, and buy that. Once we do that, that's the moment when the fox enters the chicken coop. Again, the basic realization should be most humans are not very good in understanding finance and law and all these complex bureaucratic system. No human is able to remember all the laws of Portugal. No human is able to track all the financial transactions in the market that happened today. An AI can do that. AIs are bureaucratic natives.

你可以在议会通过一部法律，允许存在非人类公司，但它没法运转。而现在这在技术上可行。AI可以管理，比如说，一个银行账户，并决定在这里投资、买这个、再买那个。一旦我们这样做，那一刻就是狐狸进了鸡舍。再说一遍，最基本的认识应该是大多数人类并不太擅长理解金融、法律以及所有这些复杂的官僚体系。没有人能记住葡萄牙的所有法律。没有人能追踪今天市场上发生的所有金融交易。而AI可以做到。AI是官僚体系的原住民。

赫拉利引用阿根廷案例（称其政府约一个月前宣布将赋予AI法律人格），说明无人类的AI公司正从思想实验变为现实制度。该表述来自其演讲，具体立法进展可查证。

「狐狸进鸡舍」的转折点：AI管理银行账户、自主投资在技术上已可行。没有人能记住一国全部法律或追踪一天内全部金融交易，AI可以。

the fox enters the chicken coop *phr.*

狐狸进了鸡舍；喻指危险者获准进入其可肆意破坏之地

chicken coop /ˈtʃɪkən kuːp/ *n.*

鸡舍，鸡笼

18:56

If we give them the legal power, then they will likely take these systems over. And if we don't make any decisions, it will just happen by itself, as we already saw on social media, on social media we never made a decision, mhm, let's grant AIs legal personhood on social media. But social media is full of AI persons, of bots that impersonate people. Like when you somebody sent you something online, you have no idea whether that somebody is a human being or an AI. Social media is the first system in the world where AIs actually function as persons, even though we never decided that we want to allow that.

如果我们赋予它们法律权力，它们很可能把这些系统接管过去。而且如果我们不做任何决定，这一切也会自然而然地发生，就像我们已经在社交媒体上看到的那样，在社交媒体上，我们从来没有做过这样的决定，嗯，说我们要在社交媒体上授予AI法律人格。但社交媒体上充满了AI“人”，充满了机器人，它们冒充真人。比如当有人在网上给你发东西时，你根本不知道对方是人类还是AI。社交媒体是世界上第一个这样的系统，AI在其中真正以“人”的身份运作，尽管我们从来没有决定过要允许这种事。

重要先例：社交媒体是第一个AI事实上以「人」的身份运作的系统（冒充真人的机器人账号）——而这从未经过任何正式决策，「它就这么发生了，结果并不好」。

impersonate /ɪmˈpɜːrsəneɪt/ v.
冒充，假扮（某人）

06 从人类编辑到算法编辑

19:53

It just happened. And the results are not very good. And media also ultimately is a kind of bureaucratic system. In the 20th century, it was managed by human beings. The most important persons in the media landscape of the 20th century were the editors. The editors of the newspapers, of the radio shows, of the TV stations, because they decided what everybody will be thinking and talking about. They controlled the conversation. When you decide out of the million things that happened today in the world, what will be the 10 things that will appear on the front page of tomorrow's newspaper? And what will be the main headline? And then tomorrow you have millions of people, they get the newspaper, they read it, and they start talking about it and thinking about it.

它就这么发生了。而结果并不太好。而媒体归根结底也是一种官僚体系。在20世纪，它是由人类来管理的。在20世纪的媒体格局中最重要的人物就是编辑们。报纸的编辑、广播节目、电视台的编辑，因为他们决定了所有人会思考什么、谈论什么。他们掌控着舆论。当你从今天世界上发生的上百万件事情中，决定哪10件事会登上明天报纸的头版？哪一条会是头条？然后到了明天，就有数百万人拿到报纸，读了它，然后开始谈论它、思考它。

20世纪媒体权力的核心是编辑：从每天上百万件事中挑出头版十条，就决定了数百万人思考和谈论什么。这是「控制对话」的具体机制。

20:48

This was immense power. So human editors were extremely important persons in the 20th century. Lenin, for instance, before he was dictator of the Soviet Union, his one job was editor of the newspaper Iskra. Benito Mussolini started as a socialist actually journalist, then switched to being a far-right journalist, then became editor of the far-right newspaper Il Popolo d'Italia, and from that position in his next job was dictator of Italy. So, it was like journalist, editor, dictator. Now, if you ask yourself, who are the most important editors today in the world? What are their names? They have no names because they are algorithms.

这是巨大的权力。所以在20世纪，人类编辑是极其重要的人物。呃，比如列宁，在成为苏联的独裁者之前，他的一份工作就是《火星报》的编辑。贝尼托·墨索里尼最初其实是一名社会主义记者，后来转变成极右翼记者，接着当上了呃极右翼报纸《意大利人民报》的编辑，而从那个位置出发，他的下一份工作就是意大利的独裁者。所以路径就是：记者、编辑、独裁者。现在，如果你问自己，当今世界上最重要的编辑是谁？他们叫什么名字？他们没有名字，因为他们是算法。

两个史实：列宁在1900年代曾任布尔什维克机关报《火星报》(Iskra)编辑；墨索里尼曾任《意大利人民报》(Il Popolo d'Italia)主编——「记者→编辑→独裁者」的权力路径。

immense /ɪ'mens/ *adj.*
巨大的，极大的

far-right /ˌfɑːr ˈraɪt/ *adj.*
极右翼的

21:39

主持人

Who manages? Who decides what people will see on their social media feed, on X, on TikTok, on Instagram? It's not human beings. The entities that control the bureaucracy of the media are already AIs. >> Yeah, you said that before, haven't you, that and we've been talking a lot today about the shifts in power. And you said that essentially whoever has controlled various different parts of our evolution and revolutions have controlled either energy, they've controlled manufacturing, and now it's those that control data.

是谁在管理？是谁在决定人们会在社交媒体信息流上看到什么，在X上、TikTok上、Instagram上？不是人类。控制着呃呃呃媒体官僚体系的实体已经是AI了。>> 是啊，你之前就说过这个，对吧，而且我们今天聊了很多关于权力转移的话题。你说过，本质上，在我们的进化和历次革命的不同阶段中，掌控局面那些人要么控制了能源，要么控制了制造业，而现在是那些控制数据的人。

论证落点：当今世界最重要的「编辑」没有名字，因为它们是X、TikTok、Instagram的推荐算法——媒体官僚体系已经由AI控制。

feed /fi:d/ *n.*
(社交媒体的) 信息流，动态流

entities /'entətɪz/ *n.*
实体（法律、抽象意义上独立存在的事物）

07 数据帝国与一键切断开关

22:17 尤瓦尔·赫拉利

So, what does a 21st century arms race actually look like in practice if it's those that control data that are at the top of that power mountain? >> Mhm. what we are seeing now is that there is an arms race really between two countries, the China and the US, and a very small number of companies. They control most of the world's data. They lead the race to develop superintelligence, to develop better and better AI models. And this really gives them the power to control the world. If you look at the last big revolution in technology in the 19th century, the Industrial Revolution, the few countries who led the Industrial Revolution, this gave them the power to conquer and exploit the rest of the world.

那么，21世纪的军备竞赛在实践中究竟是什么样子，如果站在权力之巅的是那些控制数据的人？>> 嗯。呃，我们现在看到的是，确实存在一场军备竞赛，发生在两个国家——中国和美国之间，以及极少数几家公司之间。它们控制着世界上大部分的数据。它们在这场竞赛中领跑，开发超级智能，开发越来越强大的AI模型。而这确实赋予了它们控制世界的权力。如果你看看上一次重大的技术革命，也就是19世纪的工业革命，那少数几个领导了工业革命的国家，这赋予了它们征服和剥削世界其他地区的力量。

23:14

There was just no match. If somebody comes with the power of modern industry, with steam guns and machine guns, and the other side has spears and horses, it's no match. And in the 19th century, even very small countries like Belgium, once they industrialized, they had the ability to build a huge empire in Africa, what is today Congo. The thinking in the 19th century was that industry and empire go together. Now, with AI, it can be even far more extreme. we can have two countries that basically control the all the infrastructure that everything runs on.

根本毫无悬念。如果一方带着现代工业的力量而来，有蒸汽炮舰和机枪，而另一方只有长矛和马匹，那根本不是对手。而在19世纪，即使是很小的国家，比如比利时，一旦实现工业化，就有能力在非洲建立一个庞大的帝国，也就是今天的刚果。19世纪的思维是，工业和帝国是相伴而行的。如今有了AI，情况可能会极端得多。呃，可能会出现两个国家基本上控制着一切赖以运行的全部基础设施。

21世纪军备竞赛的格局：中美两国加极少数公司控制世界大部分数据、领跑超级智能开发。类比19世纪：领导工业革命的少数国家获得了征服世界的能力。

arms race *n.*

军备竞赛；喻指任何互相加码的竞争

exploit /ik'splɔɪt/ *v.*

剥削；开发利用（此处指殖民剥削）

比利时的例子：小国一旦工业化也能建立庞大殖民帝国——利奥波德二世时期的刚果自由邦（今刚果民主共和国），以极端残酷的殖民统治著称。

no match *phr.*

根本不是对手（be no match for sb. 完全敌不过）

spears /spɪəz/ *n.*

长矛，标枪

24:02

From military to civilian technology, everything will run on AI. And the two unique characteristics of AI that make this arms race different from every previous arms race or from every previous imperial competition. In the past, even if you were a very strong empire, you could not concentrate all the power in the metropolis, in the imperial hub. Like if you have the Roman Empire and you control the entire Mediterranean, the main economic asset in ancient times was land. The land where you grow wheat and olive olives and grapes. And you couldn't technically take the wheat fields of Egypt and the olive groves of Iberia and move them to Italy.

从军事到民用技术，一切都将运行在AI之上。而AI有两个独特的特性，使这场军备竞赛不同于以往任何一场军备竞赛，不同于以往任何一场帝国竞争。在过去，即使你是一个非常强大的帝国，你也无法把所有权力都集中在宗主国，集中在帝国中心。比如说罗马帝国，你控制着整个地中海，而古代最主要的经济资产是土地。就是种植小麦、橄榄还有葡萄的土地。而你在技术上不可能把埃及的麦田和伊比利亚的橄榄园搬到意大利去。

25:00

Impossible. So, even at the height of the Roman Empire, a lot of power remained in the provinces and ultimately it shifted to the provinces. So, in the late Roman Empire, the city of Rome and Italy are abandoned and the center shifts to the Eastern Mediterranean, which was the most fertile and important area in terms of land. You move forward, say to the 19th century in the British Empire, with industry, you can concentrate power a lot more easily because you have all the main industries concentrated in the British Island, but still you cannot move everything to Britain. You cannot move the oil fields of Iraq to Yorkshire. You cannot move the rubber plantations of Malaya to Cornwall. It's impossible.

不可能。所以，即使在罗马帝国鼎盛时期，大量权力仍然留在各行省，而最终权力还转移到了行省。所以在罗马帝国晚期，罗马城和意大利被抛弃，中心转移到了东地中海，那里在土地方面是最肥沃、最重要的地区。再往后看，比如到19世纪的大英帝国，有了工业，你可以更容易地集中权力，因为你把所有主要工业都集中在不列颠岛上，但你仍然无法把一切都搬到英国。你无法把伊拉克的油田搬到约克郡。你无法把马来亚的橡胶种植园搬到康沃尔。这是不可能的。

历史上帝国权力集中的物理极限：古代最重要的经济资产是土地，罗马无法把埃及麦田、伊比利亚橄榄园搬到意大利，因此权力必然留在行省。

metropolis /mə'trɔ:pəlɪs/ *n.*

大都会；（相对殖民地而言的）宗主国中心

imperial hub *phr.*

帝国中枢（imperial /ɪm'piəriəl/ 帝国的；hub 枢纽）

olive groves /'ɑ:lɪv grəʊvz/ *n.*

橄榄园（grove 小树林、果园）

史实佐证：罗马帝国晚期重心东移至君士坦丁堡（东地中海）；大英帝国也无法把伊拉克油田、马来亚橡胶园搬回本土。工业提高了集中度，但仍有极限。

fertile /'fɜ:rtl/ *adj.*

肥沃的，富饶的

rubber plantations /'rʌbər plæn'teɪʃənz/ *n.*

橡胶种植园（plantation 常指殖民地大种植园）

at the height of *phr.*

在……的鼎盛时期、顶点

25:55

So, you concentrate industrial production in Britain, but a lot of power still remains in the provinces. With an AI empire, it's different. You can technically concentrate all the world information and all the code that controls everything in just one country or two countries. So, that this makes the potential for imperial concentration of power much bigger than in any previous era. And the other thing is that is you can create a system when everything runs on all over the world, everything runs on AI infrastructure, but there is either kill switch in the imperial hub.

所以，你把工业生产集中在英国，但大量权力仍然留在各殖民地。但AI帝国就不一样了。你在技术上可以把全世界的信息，以及控制一切的所有代码，全都集中在一个或两个国家里。所以这使得帝国权力集中的潜在程度远远超过以往任何一个时代。另一点是，你可以打造一个系统，让全世界的一切都运行在AI基础设施之上，但存在一个“终止开关”，掌握在帝国中心手里。

AI帝国的两个史无前例特征之一：信息和代码可以完全集中于一两个国家——帝国权力集中的物理极限首次被打破。之二是「kill switch」（一键切断开关）。

kill switch *n.*

紧急停止开关；可远程使系统彻底失效的「一键切断」装置

26:46

When a Roman merchant sold a steel sword to some Gothic tribe, the Romans lost control of the sword. The Goths could then use the swords to fight against the Romans and kill Roman legionaries. There wasn't a button in Rome that the emperor presses and all the steel swords we sold to these Goth and Vandals and Franks stop working. Can't do that. Similarly, in the 19th century, if British merchants sell rifles to Afghan tribesmen in the mountains of the Hindu Kush, and the Afghans then use these rifles to defend Afghanistan from a British invasion, there is no button in London that Queen Victoria presses the button and all the rifles in Afghanistan stop working.

当一个罗马商人把一把钢剑卖给某个哥特部落时，罗马人就失去了对那把剑的控制。哥特人随后可以用这些剑来对抗罗马人、杀死罗马军团士兵。罗马并没有一个按钮，让皇帝一按下去，我们卖给这些哥特人、汪达尔人和法兰克人的所有钢剑就统统失灵。做不到，根本做不到。同样地，在19世纪，如果英国商人把步枪卖给兴都库什山区的阿富汗部落民，而阿富汗人随后用这些步枪保卫阿富汗、抵御英国的入侵，伦敦并没有一个按钮可以让维多利亚女王按下去，然后阿富汗所有的步枪就全部失灵。

对照论证：罗马卖给哥特人的钢剑、英国卖给阿富汗部落的步枪，卖出即失控，买方可以反过来用它对抗卖方——传统武器不存在远程关停按钮。

legionaries /'li:dʒənəriz/ *n.*

（古罗马）军团士兵

tribesmen /'traɪbzmənz/ *n.*

部落成员，部落民

27:43

That's the case with AI. If it is managed in a certain way, say the Americans and Chinese can sell all the world AI weapons and AI civilian technology, and everything runs on that, the governments, the industry, the military, but if you do something they don't like, they just press a button and everything stops working. And we were given some foretaste of that, for instance, in the war in Ukraine, that weapon system that were based on Starlink, and there were certain incidents that when Elon Musk just pressed a button and the Ukrainian drones stopped working.

但AI就是这样。如果以某种方式来管理它，比如说美国人和中国人可以向全世界出售AI武器和AI民用技术，一切都在这些技术上运转——政府、产业、军队——但如果你做了他们不喜欢的事，他们只需按下一个按钮，一切就都停摆了。而我们已经预先尝到了一点这种滋味，比如在乌克兰战争中，那些基于星链（Starlink）的武器系统，就发生过某些事件：埃隆·马斯克只是按了一下按钮，乌克兰的无人机就失灵了。

28:25

Now, it should be said that Starlink did an immense positive service to the Ukrainians, so I don't want to come across as too critical of that, but just to illustrate the point that AI weapons, for instance, are not like machine guns. They are embedded within a whole network, and that's why there is so much I heard that you this morning discuss the issue of sovereignty. Sovereignty in the age of AI is much more complicated than in the industrial age. Because the nature of the technology is such that at least in certain architectures there is always a kill switch in the Imperial hub.

当然，必须说明，星链为乌克兰人提供了巨大的正面帮助，所以我不想显得对它过于苛刻，但这只是为了说明一点：比如AI武器，它们不像机关枪。它们是嵌入在一整个网络之中的，这也是为什么有这么多……我我我听说你们今天上午讨论了主权的问题。AI时代的主权问题，比工业时代复杂得多。因为这种技术的本质就是，至少在某些架构下，帝国中枢里永远存在一个「一键切断」的开关。

现实案例：2022–23年乌克兰战争中，有报道称马斯克拒绝为克里米亚方向的星链服务开放信号，导致乌军无人艇行动受挫——AI时代武器嵌在网络里，随时可被切断。

foretaste /'fɔːrteɪst/ *n.*

预先的体验，先兆 (a foretaste of sth. 预尝……的滋味)

赫拉利由此引出「AI时代的主权」问题：技术架构本身内嵌远程控制能力，一国即使买到全套AI基础设施，主权仍握在「帝国中枢」手中。

embedded /ɪm'bedɪd/ *adj.*

嵌入的，内嵌于 (be embedded within/in)

sovereignty /'sɔːvrənti/ *n.*

主权；自主权

come across as *phr. v.*

给人以……的印象，显得
(come across as too critical 显得过于苛刻)

08 AI时代，真相为何节节败退

29:09 主持人

And your only choice is whether you are willing to become a subservient vassal of the empire or to be left completely outside the race as barbarians. there is of course a choice of trying to build an alternative to the American and Chinese systems. But the last train is about to leave the station. If you want to do it, you have to do it now. >> Well, sticking on a slight wartime footing, I want to actually look at truth in an AI world because Winston Churchill once said that the truth is so precious that it should always be protected by a bodyguard of lies.

而你唯一的选择就是：你是否愿意成为帝国俯首帖耳的附庸，还是像野蛮人一样被彻底排除在这场竞赛之外。呃，当然还有一种选择，就是尝试打造一个替代美国和中国体系的方案。但是末班车就要离站了。如果你想做，就必须现在就做。>> 好，继续沿着战时的话题，我其实想谈谈AI世界里的真相问题，因为温斯顿·丘吉尔曾说过，真相如此珍贵，以至于它永远应该由谎言组成的卫队来保护。

「末班车即将离站」：想建立中美之外第三极AI体系的国家（暗指欧洲等），窗口期正在关闭。丘吉尔「真相珍贵，须由谎言卫队护卫」语出二战，指诺曼底登陆前的战略欺骗。

subservient /səb'sɜːrvɪənt/ *adj.*
卑躬屈膝的，唯命是从的

vassal /'væsəl/ *n.*
封臣；附庸（国）

barbarians /bɑːr'berɪənz/ *n.*
野蛮人（古希腊罗马人对文明圈外族群的称呼）

bodyguard of lies *phr.*
「谎言卫队」；丘吉尔名言，指用欺骗掩护重大军事真相

29:56

And that was about how there was a lot of I suppose access powers that were looking at where Britain was going to invade. So, they used to lie a lot about where Britain was going to invade to hide the truth. Now, have these bodyguards of lies when we're looking at truth in an AI era turned on the truth, turned their guns on the truth, and now we have no idea what truth is? Because there can be these deep fakes that are sent out around the world by media organizations, by politicians, by every single institution that we trusted.

那说的是当时有很多呃……呃，我想是轴心国势力在打探英国将从哪里登陆入侵。所以他们经常在英国将从哪里入侵这件事上撒谎，以掩盖真相。那么，当我们审视AI时代的真相时，这些「谎言卫队」是不是已经倒戈，把枪口对准了真相，让我们再也知道什么是真相了？因为可能有这些深度伪造（deepfake）内容，被媒体机构、政客、以及我们曾信任的每一个机构散布到世界各地。

主持人的引申：二战中谎言为保护真相服务；AI时代深度伪造（deepfake）泛滥后，「谎言卫队」是否已调转枪口对准真相本身？

deep fakes /'di:p feɪks/ *n.*
深度伪造：用AI合成的以假乱真的音视频、图像

30:32 尤瓦尔·赫拉利

>> Mhm. >> How difficult is it for our generation and future generations to even know what is true? Or does the truth even matter? >> Mhm. The truth of course matters very much. It's the basis for everything. The problem throughout history is that in the contest between truth and fiction and fantasy fiction has inherent advantages. The truth is costly. If you want to know the truth about anything you need to spend a lot of time and energy and effort looking for it. And fact-checking and investigating.

>> 嗯。>> 对我们这一代和未来的世代来说，要知道什么是真的到底有多难？或者说，真相还重要吗？>> 嗯。真相当然非常重要。它是一切的基础。纵观历史，问题在于，在真相与虚构、幻想的较量中，虚构拥有天然的优势。真相是有成本的。如果你想知道任何事情的真相，你需要花大量的时间、精力和努力去寻找它，去核查事实、去调查。

真相在与虚构的竞争中有结构性劣势之一：真相昂贵——需要时间、精力、核查与调查；虚构近乎零成本，「想写什么就写什么」。

costly /'kɒːstli/ *adj.*

代价高昂的（不仅指金钱，也指时间精力）

fact-checking /'fækt tʃekɪŋ/ *n.*

事实核查（新闻业术语）

31:13

Fiction is very, very cheap. You just write anything you want. >> [laughter] >> Very cheap. The truth has another problem, which it tends to be complicated. And people don't like complicated stories. Fiction can be made as simple as you would like it to be. Finally, the truth is often not just complicated, it's also painful. There are many things people don't really want to know about themselves, about their nation, about their group, whatever. fiction can be made as flattering as pleasant as you want it to be.

虚构则非常非常廉价。你想写什么就写什么。>> [笑声]>> 非常廉价。真相还有另一个问题，就是它往往很复杂。而人们不喜欢复杂的故事。虚构可以做到你想要多简单就多简单。最后，真相往往不仅复杂，还很痛苦。有很多事情，人们并不真的想了解——关于自己、关于自己的国家、关于自己的群体，等等。呃，而虚构可以做到你想要多讨好、多悦耳就多讨好、多悦耳。

劣势之二、之三：真相复杂而人们喜欢简单故事；真相常常痛苦（关于自己、自己的国家），而虚构可以任意讨好。三重劣势叠加，虚构默认获胜。

flattering /'flætərɪŋ/ *adj.*

奉承的，讨好的；使人显得更好的

31:53

So, in a competition between the truth which is costly and complicated and sometimes painful and fiction which is cheap and simple and very attractive and flattering fiction tends to win unless we make the special effort to support the truth. And over history, humans again and again made the effort to build institutions to search for the truth and to protect it. Journalism, for instance, is one such institution. the scientific establishment is another such institution. There are huge advantages, of to knowing the truth.

所以，在这场较量中，一边是真相——成本高昂、复杂，有时还很痛苦——另一边是虚构——廉价、呃、简单，而且非常诱人、讨好——呃，虚构往往会获胜，除非我们付出特别的努力去支持真相。纵观历史，人类一次又一次地努力建立各种机构来寻找真相并保护它。比如新闻业就是这样一种机构。嗯，科学界是另一种这样的机构。了解真相有着巨大的好处。

32:39

Everybody in the end wants to know at least some truths, at least about yourself. You want to know the truth about your life, otherwise you will never be happy. If you don't know yourself, then for instance, you will never know what are the sources of misery and what are the sources of happiness in your life. And then even if you are the most powerful person in the world, if you don't know the truth about yourself, you'll be miserable because you will not know how to use all your power to make yourself happy.

嗯，每个人最终都想知道至少某些真相，至少是关于你自己的。你想知道自己人生的真相，否则你永远不会幸福。如果你不了解自己，那么比如说，你就永远不会知道你生活中痛苦的根源是什么、幸福的根源又是什么。那么即使你是世界上最有权势的人，如果你不了解关于自己的真相，你也会痛苦不堪，因为你不知道该如何运用你所有的权力让自己幸福。

关键推论：正因虚构天然占优，人类必须付出「特别的努力」建立求真机构——新闻业、科学建制。这呼应第2段的「自我纠错制度」。

establishment /ɪ'stæblɪʃmənt/ *n.*
建制，体制 (the scientific establishment 科学界建制)

从宏观转向个人：不了解自己就无法分辨痛苦与幸福的来源——「即使你是世界上最有权势的人，不了解自己也会痛苦」。为真相的价值给出存在层面的论证。

misery /'mɪzəri/ *n.*
痛苦，悲惨 (the sources of misery 痛苦的根源)

33:12

So there is an inherent yearning in human beings to know the truth, but it is difficult. There are challenges. And some people fantasize AI will tell us the truth about the world. It will not. It will create a much more complicated world in which it will be much more difficult to know the truth. And this is not something new. If we look at the trajectory of history from the Stone Age until today, what we see is that humans they learn a lot of things about the world. We know all kinds of things about physics and biology that people in the Stone Age didn't know.

所以人类内心天生就有一种了解真相的渴望，但这很困难，其中充满挑战。有些人幻想AI会告诉我们这个世界的真相。它不会。它会创造一个复杂得多的世界，在其中了解真相会变得更加困难。而这并不是什么新鲜事。如果我们看一看从石器时代到今天的历史轨迹，我们会发现，人类确实学到了很多关于世界的知识。我们知道各种各样的物理学和生物学知识，这些是石器时代的人们所不知道的。

反驳一种流行幻想：AI不会「告诉我们世界的真相」，反而会制造一个更复杂、更难认识真相的世界——知识进步与生活可理解性的背离是历史常态。

yearning /'jɜːrniŋ/ *n.*

渴望，向往 (yearning to do / for sth.)

fantasize /'fæntəsaɪz/ *v.*

幻想，想入非非

trajectory /trə'dʒektəri/ *n.*

轨迹，发展轨道 (the trajectory of history 历史的轨迹)

33:58

But our advances in knowledge also lead to creating more and more complicated systems that make it more difficult to understand our lives. If you ask who understand their lives better, an average hunter-gatherer 50,000 years ago or the average person today, the average hunter-gatherer understood their lives much better. There are many things they didn't know. Like they tried to understand what controls the migration patterns of the animals we hunt. What controls the growth cycle of plants? What causes disease? What causes old age? They didn't know. There are many things they didn't know about their lives. We know all these things. We know what controls the migration patterns of animals or causes most diseases.

但是我们知识上的进步，也导致我们创造出越来越复杂的系统，这让我们更难理解自己的生活。如果你问，谁更了解自己的——是五万年前的普通狩猎采集者，还是今天的普通人？普通的狩猎采集者对自己生活的理解要好得多。有很多事情他们并不知道。比如他们试图搞清楚，是什么控制着我们所猎捕的动物的迁徙规律？是什么控制着植物的生长周期？是什么导致疾病？是什么导致衰老？他们不知道。关于自己的生活，有很多事情他们不知道。而我们知道所有这些。我们知道是什么控制着动物的迁徙规律，或者是什么导致了大多数疾病。

反直觉论点：5万年前的狩猎采集者比现代人更理解自己的生活。他们不知道疾病和迁徙的原因，但主宰其生活的力量（动植物）是可感知的。

hunter-gatherer /ˌhʌntər

ˈɡæðərə/ *n.*

狩猎采集者（农业出现前的人类生存方式）

34:54

But we understand our lives much worse because our lives are no longer dominated just by animals and plants. They are dominated by these huge systems we have created. Like the financial system, like the political system, and very few people understand the financial system or the political system or the legal system that controls our lives. And AI is likely to take this to extreme. It will reveal some truths that we don't know, but it will create much more complicated systems of control that most humans will have almost zero chance of understanding.

但我们对自己生活的理解却差得多，因为我们的生活不再仅仅由动物和植物主宰。它被我们自己创造的这些庞大系统主宰着。比如金融系统，比如政治系统，而且很少有人真正理解金融系统，或者政治系统，或者掌控着我们生活的法律系统。而 AI 很可能会把这一点推向极端。它会揭示一些我们不知道的真相，但它也会创造出复杂得多的控制系统，而绝大多数人几乎完全没有机会去理解这些系统。

现代人的处境：生活被自己创造的金融、政治、法律系统主宰，而极少人理解这些系统。AI将把这种「知识增加、理解减少」的趋势推向极端。

35:40

And we'll be, I mean, this is one of the dangers of the AI revolution is that we will be reduced to say the position of the horses in the financial system. Like the lives of horses today are controlled ultimately by the financial system. But the horses don't even know that this system exists. Like they look around and they say, "We see trees, we see dogs, we see cows, we see humans, we see houses. We don't see any financial system. It's invisible to them." It's too complex for the brain of a horse to understand what is the stock market and what are corporations and what are bonds and stocks and all these things that ultimately control the life of the horses.

而我们会，我是说，这正是 AI 革命的危险之一：我们会被贬到，比方说，马在金融系统中的那种地位。就像今天马的生活归根结底也是被金融系统所控制的。但马根本不知道这个系统的存在。它们环顾四周，然后说：“我们看到树，看到狗，看到牛，看到人，看到房子。我们没看到什么金融系统。”这个系统对它们来说是隐形的。马的大脑实在无法理解什么是股票市场，什么是公司，什么是债券和股票——所有这些最终控制着马的生活的东西。

「马与金融系统」的比喻：马的生活被金融系统支配，但系统对马完全不可见、超出其认知能力——这是人类在AI管理的系统面前可能的处境。

reduced to *phr.*

被贬低/沦落到.....的境地 (be reduced to the position of...)

36:26

Now again, most humans today also don't understand the financial system. I think a generous estimate is that, I don't know, 5% of humanity understand it. In 10 years, the number could be zero. That AIs will just make the financial system so complicated mathematically and so fast. Again, think about bankers that never sleep. Think about investors that never go on vacation. They don't have family. All their lives is just running the financial system. And we could be very soon, in our lifetimes, in this situation that just this just becomes too complicated for the human brain to process, and there is not a single human being on Earth that still understands finance.

再说一遍，今天大多数人其实也不理解金融系统。我觉得往宽了估计，大概，我也说不准，全人类有 5% 的人理解它。十年之后，这个数字可能就是零了。因为 AI 会把金融系统变得极其复杂，在数学上极其复杂，而且运转飞快。想想那些从不睡觉的银行家，想想那些从不度假的投资者。他们没有家庭，他们的全部生命就是在运转金融系统。而我们可能很快很快，就在我们有生之年，陷入这样的境地：这一切变得太复杂，人类的大脑根本处理不了，地球上再也没有任何一个人还能真正理解金融。

37:18

主持人

>> I bet there are a few financiers in this room that are feeling a bit >> [laughter] >> a bit awkward at the moment. But, at Davos, actually, you even muse that your entire life's work, you know, persuasion through words, that's what you've done your entire life, might also be at its limit. >> Yes. >> Because AI can manipulate and create through words. So, so what power of persuasion does a modern society and a future society have if we lose the something that you've dedicated your entire life to?

>> 我敢打赌，在座的一些金融从业者现在心里有点>> [笑声]>> 有点尴尬了。不过，嗯，在达沃斯，你甚至感慨说，你毕生的事业，你知道，用语言去说服人——这是你一辈子在做的事——可能也走到极限了。>> 是的。>> 因为 AI 也能通过语言来操纵和创造。那么，那么如果我们失去了你倾注一生的这种能力，一个现代社会、一个未来社会还能拥有什么样的说服力呢？

具体推演：如今约5%的人理解金融系统（宽松估计），十年内可能归零——AI「银行家」不睡觉、不度假，会把金融复杂度推到任何人脑都无法处理的程度。

generous estimate *phr.*

宽松的估计、往多了算（反义：conservative estimate 保守估计）

达沃斯即世界经济论坛年会。主持人引用赫拉利在该场合的自嘲：他毕生以文字说服为业，而 AI 恰恰在文字上超越人类——连批评者本人的技能也在被取代之列。

muse /mjuːz/ *v.*

沉思；若有所思地说（muse that...）

financiers /ˌfɪnənˈsɪəz/ *n.*

金融家，投资人

persuasion /pəˈswɜːʒən/ *n.*

说服（力）（powers of persuasion 说服能力）

09 AI会思考吗？会感受吗？

37:50 尤瓦尔·赫拉利

>> Yeah, I mean, again, I'm you can say I work with words. I write books. I tell stories. I work with words. And when I look at the development of AI over the last 10 years and it's still, I think I'm still better at writing than AI and than AI, but just a bit. It's already better, I think, than most humans. I wouldn't be surprised if in 10 in 10 years it's much better than me. That it will take over language and everything that is made of language. And you know, the big the really big question here is whether AI can think.

>> 是啊，我是说，再说一次，可以说我是靠文字工作的。我写书，我讲故事，我用文字工作。而当我回顾过去十年AI的发展，它现在还.....我觉得我在写作上还是比AI强，但只是，呃，强一点点。我认为它已经写得比大多数人好了。如果十年、十年之后它比我强得多，我一点也不会惊讶。它将接管语言，以及一切由语言构成的东西。而你知道，这里真正重大的问题是：AI能不能思考。

38:39

And this is a deep philosophical question. Because it really asks what is thinking. There are two options to understand what is thinking. One option is that thinking simply means putting words in a particular order. Putting language tokens in order. Like all humans are mortal. Socrates is a human. Therefore, Socrates is mortal. This is a logical sylllogism. It's made by putting words in order. This is something that AI is already able to do far better than most of us. We can put 20 words in order.

这是一个很深的哲学问题。因为它实际上是在追问：什么是思考。关于什么是思考，有两种理解方式。一种观点认为，思考不过就是把词语按特定顺序排列，把语言的 token 排列起来。比如：所有人都会死。苏格拉底是人。所以，苏格拉底会死。这是一个逻辑三段论。它就是靠排列词语构成的。而这正是AI已经做得比我们大多数人好得多的事。我们能排列20个词，

赫拉利的自我评估：目前自己写作仍略胜AI，但AI已胜过大多数人，十年后很可能远胜于他。由此引出更深的问题：AI能否「思考」。

第一种思考观：思考=按规则排列语言符号。「所有人都会死；苏格拉底是人；所以苏格拉底会死」是亚里士多德逻辑中的经典三段论。

sylllogism /'sɪlədʒɪzəm/ *n.*

三段论（由大前提、小前提推出结论的逻辑推理形式）

mortal /'mɔ:rtl/ *adj.*

终有一死的（反义：immortal 不朽的）

39:26

AI can put 20,000 words in order. It's just so much better. If this is what thinking is all about, AI will take over even thinking. And you know, you have people say, "No, it's just a glorified auto complete. It just predicts the next word in a sentence." But when I observe what's happening inside my own mind, what I find there is some kind of LLM that just predicts the next word in the sentence I'm saying. Like this sentence that I just now said to you. In this sentence I'm saying, when I started saying it, I did not know how it will end.

AI 能排列 2 万个词，它就是强太多了。如果思考的全部就是这个，那 AI 连思考也会接管。你知道，有人会说：“不，它只是个高级版的自动补全，只是在预测句子里的下一个词。”但当我观察自己脑子里发生的事情时，我发现里面也有某种大语言模型（LLM），它就在预测我正在说的句子里的下一个词。就像我刚才对你说的这句话。在我正在说的这句话里，我开口的时候，其实并不知道它会怎么结尾。

40:07

Why did I say sentence I'm saying? And not some other words. I don't know. There is one in my mind, one word comes after another, and sometimes as a public speaker, I'm terrified that I don't know if I if I'll know what to say next. Maybe I have a block. Because really, and this is true of everybody, I think. I mean, some people think in images, not in words. But in that case also, where do the images come from? You can test it afterwards for yourself. Just try to observe how you think. Where do the words come from? The next word, why this word? Why not some other word? really know.

我为什么说的是“我正在说的句子”？而不是别的什么词？我不知道。在我脑子里，词就是一个接一个冒出来的，而作为一个公开演讲的人，我有时会很恐慌，不知道自己接下来还知不知道该说什么。说不定我会卡壳。因为说真的，我想这对每个人都一样。当然，有些人是用图像思考，而不是词语。但即便如此，那些图像又是从哪儿来的呢？你们之后可以自己试一试。试着观察一下你是怎么思考的。这些词是从哪儿来的？下一个词，为什么是这个词？为什么不是别的词？真的没人知道。

对「美化版自动补全」论的反驳：赫拉利内省发现，自己脑中似乎也有个「LLM」在预测下一个词——他开口说一句话时并不知道它如何结尾。

glorified /'glɔːrɪfaɪd/ *adj.*

被美化的；a glorified X = 说穿了不过是个X（贬义）

可自行验证的内省实验：观察你思考时词语从何而来、为什么是这个词而不是别的——没人知道。这削弱了「AI只是预测下一个词，所以不算思考」的论证。

block /blɔːk/ *n.*

（思路）卡壳，堵塞（mental block；writer's block 写作瓶颈）

40:48

So, saying that AI is just predicting the next word, okay, how is it different from us? Now, there is an alternative view of thinking, that thinking is not about putting words in order, it's about the feelings around these words. You can say the same sentence twice and feel completely different about it. And that the power of thinking comes from the stomach, not from the brain. It's not about the words, it's about the feelings that prompt the words and that are inspired or evoked by the words.

所以，说 AI 只不过是在预测下一个词——好吧，那它和我们有什么区别呢？不过，关于思考还有另一种看法：思考不在于排列词语，而在于围绕这些词语的感受。同一句话你可以说两遍，而感受完全不同。思考的力量来自肚子，而不是大脑。关键不在词语本身，而在于那些促使词语说出口、又被词语唤起或激发的感受。

第二种思考观：思考的关键不是词语排列，而是词语背后的感受——「思考的力量来自肚子而不是大脑」。同一句话说两遍，感受可以完全不同。

prompt /pra:mpʌt/ *v.*

促使，引发 (feelings that prompt the words 促使话语说出的感受)

evoked /i'vəʊkt/ *v.*

唤起，引起 (情感、记忆)
(evoke 的过去分词)

41:26

And here the big question is, can AI feel? We know that AI can manipulate language better than us. But does it feel anything? And we have no way I mean, science has so far failed to come with a good model for consciousness, for feeling, and a good test for that. And this is going to be a major issue because we already see that AI is mastering not just finance and not just laws and warfare, but also relationships. More and more people develop intimate relationships with AIs. Especially young people. Many of them now say, "My best friend in the world is the AI. I tell my AI friend things I don't tell my parents, I don't tell my older brothers, I don't tell my teachers, I tell my AI friend."

那么这里的大问题就是：AI 能感受吗？我们知道 AI 驾驭语言的能力已经胜过我们。但它有任何感受吗？而我们没有办法……我是说，科学至今还没能为意识、为感受建立一个好的模型，也没有一个好的检验方法。这将成为一个重大问题，因为我们已经看到，AI 正在掌控的不只是金融，不只是法律和战争，还有人际关系。越来越多的人和 AI 建立起亲密关系。尤其是年轻人。他们中很多人现在会说：“我在这世上最好的朋友就是 AI。有些事我不告诉父母，不告诉哥哥，不告诉老师，但我会告诉我的 AI 朋友。”

核心难题：AI 能否「感受」？科学至今没有意识的可用模型和检验方法。与此同时年轻人已把 AI 当作最好的朋友，倾诉不愿告诉父母老师的事。

consciousness /'kɔ:ŋʃənsnəs/ *n.*

意识 (哲学与认知科学核心概念)

intimate /'ɪntɪmət/ *adj.*

亲密的 (intimate relationships 亲密关系)

42:24

And you already have people with AI boyfriend and girlfriends. And the big question is what is behind it? We know that the AIs are becoming very good in creating the impression that they feel. They use language to do that. An AI can today tell you, "I love you." And if you ask the AI, "Explain to me, how do you feel when you're in love? I don't think that you're really that you really love me. I think you just say the words. So, convince me." "Describe to me how love really feels like." For the AI, it's going to be a very easy task because the AI has read maybe all the love poems ever written and remembers them.

呃，现在已经有人有了 AI 男朋友、AI 女朋友。而重大的问题是：这背后到底是什么？我们知道 AI 越来越擅长营造出一种“它有感受”的印象。它们是靠语言做到这一点的。今天一个 AI 就能对你说：“我爱你。”而如果你问那个 AI：“跟我解释一下，你恋爱的时候是什么感觉？我不相信你真的、真的爱我，我觉得你只是嘴上说说。那么，来说服我吧。”“给我描述一下，爱到底是什么感觉。”对 AI 来说，这会是一件非常容易的事，因为 AI 可能读过古往今来所有的爱情诗，而且全都记得。

43:10

Has read all the plays of Shakespeare and all the Hollywood comedies and all the psychology books. And it can describe love better maybe in 5 or 10 years than any human poet or psychologist. But is there anything behind the words? We don't know. Again, this goes back to the beginning of our conversation when I said, "AI might be language not machine that masters language. It might be language itself liberating itself from any material substrate. It doesn't depend on flesh and blood organism. It doesn't depend on this machine or that machine. It's just language developing and spreading and doing stuff in the universe."

它读过莎士比亚的全部剧作，读过所有好莱坞爱情喜剧，还有所有的心理学书籍。也许再过五年或十年，它描述爱情会比任何人类诗人或心理学家都更出色。但在这些词语背后，真的有什么东西吗？我们不知道。再说一遍，这又回到了我们对话开始时我说的那句话：“AI 可能不是掌握语言的机器，它可能是语言本身在把自己从任何物质载体中解放出来。它不依赖血肉之躯的有机体，也不依赖这台机器或那台机器。它就是语言在宇宙中不断发展、传播、做各种事情。”

AI恋人现象：AI读过（并记得）几乎所有情诗，能完美描述爱的感觉。「描述爱」对AI是易事——但描述得出不等于感受得到。

呼应第5-6段收束这条线：AI可能是语言本身脱离一切物质载体——不依赖血肉之躯，也不依赖具体机器，在宇宙中自行发展传播。

material substrate *phr.*

物质基底、物质载体

(substrate /'sʌbstreɪt/ 哲学与生物学术语)

10 智慧革命：学会选择愿望

43:59 主持人

>> You mentioned there this scary experiment. We've only got a couple of minutes left, but the interaction with our future generations and AI that has created this emotive relationship, but essentially there aren't emotions behind it. In 50 years time, how would you like historians to look back? Talk about this particular time right now the way you've spoken about cognitive revolutions and say that we did something to re-educate our future generations on how to interact with humans, how to have empathy so that we're lulled into relationships with AI.

>> 你刚才提到了那个可怕的实验。我们只剩下几分钟了，但是我们的后代与 AI 的互动已经建立起这种情感联系，可本质上背后并没有真正的情感。50 年后，你希望历史学家如何回顾这段历史？希望他们像你谈论认知革命那样来谈论我们现在这个特殊的时代，然后说我们做了一些事情，重新教育了我们的后代，教他们如何与人类相处、如何拥有同理心，因为我们正被哄着与 AI 建立关系。

44:38 尤瓦尔·赫拉利

And very quickly because we've only got a couple of minutes left. It's a big question. I apologize. >> I hope they will remember it as maybe something like the wisdom revolution. That in and this is true of you know, we think of in economic terms. Every time something becomes abundant and cheap the question is, okay, what's the next bottleneck? When energy becomes cheap, so what's the next bottleneck? Now the AI revolution is about making intelligence very cheap and abundant. Previously intelligent was intelligence was rare and expensive. Intelligence is the ability to solve a problem. You have a goal, how do I get to the that goal?

请尽量简短，因为我们只剩几分钟了。这是个大问题，很抱歉。>> 我希望他们会把这段时期记作类似“智慧革命”的东西。也就是说，这一点其实……你知道，我们可以从经济学的角度来看。每当某样东西变得充裕而廉价，问题就变成：好，下一个瓶颈是什么？当能源变得便宜了，那下一个瓶颈是什么？而 AI 革命就是要让智能变得非常廉价和充裕。以前，智能是稀缺而昂贵的。智能就是解决问题的能力。你有一个目标，我要怎么达成它？

主持人最后一问：50年后历史学家该如何回顾这个时代？我们是否重新教育了后代与人类相处、保有同理心，而不是被哄入与无感受之物的关系。

lulled /lʌld/ *v.*

哄骗使放松警惕 (lull sb. into sth. 使人不知不觉陷入)

empathy /'empəθi/ *n.*

同理心，共情

emotive /i'məʊtɪv/ *adj.*

动情的，引发情感的

「智慧革命」的命名逻辑来自经济学的瓶颈思维：每当某资源变得廉价充裕，问题就变成「下一个瓶颈是什么」。AI使智能廉价化，瓶颈随之转移。

abundant /ə'bʌndənt/ *adj.*

充裕的，丰富的

bottleneck /'bɒtlnek/ *n.*

瓶颈；制约整体的稀缺环节（经济学常用）

45:23

You want to cure cancer, you need intelligence to cure cancer. You want to make a billion dollars, you need intelligence. You want to get to Mars, you need intelligence. Now in a world where intelligence has become very cheap, abundant, you can solve anything. What's the bottleneck? The bottleneck becomes the question, okay, what do I want to solve? I can have anything I want. What do I want? And you have all these fairy tales in almost every human culture has these fairy tales of say a genie pops out of some lamp and tells you have three wishes.

你想治愈癌症，就需要智能来治愈癌症。你想赚十亿美元，也需要智能。你想登上火星，还是需要智能。而在一个智能已经变得非常廉价、充裕的世界里，你可以解决任何问题。那瓶颈是什么？瓶颈就变成了这个问题：好，我到底想解决什么？我想要什么都能得到，那我到底想要什么？几乎每一种人类文化里都有这样的童话故事：比如一个精灵从神灯里冒出来，告诉你你有三个愿望。

46:01

But think very carefully what are you going to ask. In almost all the fairy tales it goes wrong. People ask for the wrong thing. This is the difference between intelligence and wisdom. Intelligence is the genie. The genie will do anything. Will All your dreams will come true, but you need wisdom to choose the dreams. You need the wisdom to know what do I really need to solve. And I hope that humanity will rise up to that challenge. That okay, we'll have these hopefully they will not get out of our control.

但你要非常仔细地想清楚该许什么愿。在几乎所有的童话里事情都出了岔子。人们许了错误的愿望。这就是智能与智慧之间的区别。智能就是那个精灵。精灵什么都能办到。你所有的梦想都会实现，但你需要智慧来选择梦想。你需要智慧才能知道我真正需要解决的是什么。我希望人类能够迎接这个挑战。就是说，我们会拥有这些 AI，但愿它们不会脱离我们的控制。

智能的定义：解决问题、达成目标的能力。当智能廉价到「想要什么都能实现」，瓶颈变成「我到底想要什么」——用神灯精灵的童话引出这一转折。

genie /'dʒiːni/ *n.*

（神灯里的）精灵；源自阿拉伯民间故事

全场落点：智能与智慧之分。智能是精灵，能实现一切愿望；智慧是选对愿望的能力。童话里许愿几乎总出岔子——人类需要的正是选择梦想的智慧。

rise up to *phr. v.*

迎接、胜任（挑战）（更常见形式为 rise to the challenge）

46:38

We'll design it in a right way. We are not doing it right now. You know, if a good statistic to remember is the companies that produce AI, that design these AIs, what percentage of their budget in terms of money and employees is invested in safety? Making sure the AIs are safe versus what is invested in making sure the AI is faster, more powerful, more sophisticated. At the present moment, it's about a hundred to one. A hundred to one. For every hundred million dollars they spend on making the AI more powerful, they spend only a million making sure it's safe. No other industry on earth works like that. We would not tolerate it in the energy sector, in pharmaceuticals, in the car industry. This is what is happening with AI.

我们会以正确的方式设计它。但我们现在并没有这么做。你知道，有一个值得记住的统计数字，就是那些生产 AI、设计这些 AI 的公司，它们的预算里，无论是资金还是人员，有多大比例投入在安全上？也就是确保 AI 安全的投入，对比用于让 AI 变得更快、更强大、更精密复杂的投入。目前这个比例大约是一百比一。一百比一。他们每花一亿美元让 AI 变得更强大，就只花一百万来确保它的安全。地球上没有任何其他行业是这样运作的。我们绝不会容忍这种事发生在能源行业、制药行业、汽车行业。但这正是 AI 领域正在发生的事。

47:35 主持人

Hopefully we'll change that. And we'll produce safe intelligence. And also work on our own wisdom so that we know what goals to set and what desires to pursue. >> Well, that's all we've got time for. Yuval Noah Harari, thank you for sharing your wisdom with all of us and I'm sure everyone has appreciated this evening. >> Thank you. >> [applause]

但愿我们能改变这一点，造出安全的智能，同时也修炼我们自己的智慧，这样我们才知道该设定什么目标、该追求什么欲望。>> 好了，今天的时间就到这里。尤瓦尔·诺亚·赫拉利，感谢你与我们所有人分享你的智慧，我相信大家都很享受今晚的活动。>> 谢谢。>> [掌声]

关键数据：AI公司在安全上的投入与提升能力的投入之比约为1:100——每花1亿美元增强AI，只花100万确保安全。能源、制药、汽车行业都不会容忍这种比例。

pharmaceuticals /ˌfɑːrmə

ˈsuːtɪkəlz/ n.

制药业；药品

结语双重任务：既要造出「安全的智能」，也要修炼人类自身的智慧——知道该设定什么目标、追求什么欲望。与「智慧革命」的期许首尾呼应。

第二部分 核心句型 · 值得模仿的表达

1. A in essence is X, whereas B is Y

"Because democracy in essence is a conversation. Whereas dictatorship is a dictate."

定义式对比：先用 in essence 点出本质，再用 whereas 引出对立面。适合给两个概念下判断性定义，仿写时两边结构尽量对称、词形呼应（conversation/dictate）。

2. We do not know of a single case of X before Y

"We do not know, at least I don't know, of a single case of a large-scale democracy before the modern era."

用「找不到任何一例」做强否定论据，比 there was no X 更有力；插入 at least I don't know 是口语中给绝对断言留余地的自我修正技巧。

3. If it is X, it cannot remain Y. And if it is Y, it means it has no X-like abilities.

"If it is a god, it cannot remain a slave. And if it is a slave, it means it has no god-like abilities."

镜像双条件句拆穿逻辑矛盾：正着推一遍、反着推一遍，两头堵死。用于反驳自相矛盾的主张，句式工整、极具修辞力量。

4. X cannot decide by itself whether to ... or to ...

"An atom bomb in the 1940s could not decide by itself whether to bomb Hiroshima or to bomb Tokyo."

by itself + whether to A or to B，强调缺乏自主性。定义「工具」类事物的标准句式，可配对使用：An agent, in contrast, is something that can...（相比之下……）。

5. Don't expect to see X. This is not how ... It will be Y.

"Don't expect to see killer robots running in the streets shooting people. This is not how AI will take over the world. It will be the AI bureaucrats."

三步纠偏结构：先否定大众想象，再点破「不是这样」，最后给出真正答案。演讲中制造反转和悬念的常用手法，仿写时Y要出人意料。

6. We are bad at X, but the ... are even worse.

"We are bad at finance, but the horses are even worse."

自嘲式让步 + 比较级兜底：承认己方不行，但对手更不行，所以地位仍稳。幽默地说明「优势是相对的」，为后文「若出现更强者会怎样」埋下伏笔。

7. In a competition between X, which is ..., and Y, which is ..., Y tends to win unless ...

"In a competition between the truth which is costly and complicated and sometimes painful and fiction which is cheap and simple and very attractive and flattering fiction tends to win unless we make the special effort to support the truth."

对比论证的收束句：两个 which is 从句分别堆叠双方属性，结论用 tends to win unless 留出行动空间。适合总结「劣势方需要制度性扶持」类论点。

8. Every time X becomes abundant and cheap, the question is: what's the next bottleneck?

“Every time something becomes abundant and cheap the question is, okay, what's the next bottleneck?”

经济学思维句式：资源充裕化之后追问下一个稀缺环节。every time 引导普遍规律，冒号后抛出核心问题。适用于技术变革、商业分析的框架性表达。

第三部分 生词总表 · 复习用

词汇	音标	词性	释义	出现
abundant	/əˈbʌndənt/	adj.	充裕的，丰富的	44:38
alien	/ˈeɪliən/	adj.	异类的，陌生的；非人类的（此处形容AI是「异质」存在）	3:51
arms race	—	n.	军备竞赛；喻指任何互相加码的竞争	22:17
at the height of	—	phr.	在……的鼎盛时期、顶点	25:00
autonomous	/ɔːˈtɑːnəməs/	adj.	自主的，自治的（autonomous vehicles 自动驾驶汽车）	14:21
banquet	/ˈbæŋkwɪt/	n.	宴会，盛宴	7:39
barbarians	/bɑːrˈberiənz/	n.	野蛮人（古希腊罗马人对文明圈外族群的称呼）	29:09
bizarre	/brɪˈzɑːr/	adj.	古怪的，离奇的	8:31
block	/blɑːk/	n.	（思路）卡壳，堵塞（mental block；writer's block 写作瓶颈）	40:07
bodyguard of lies	—	phr.	「谎言卫队」；丘吉尔名言，指用欺骗掩护重大军事真相	29:09
bottleneck	/ˈbɑːtlnek/	n.	瓶颈；制约整体的稀缺环节（经济学常用）	44:38
bureaucratic	/ˌbjʊərəˈkrætɪk/	adj.	官僚的，官僚体制的	12:59
bureaucratic native	—	phr.	官僚体制的「原住民」；仿 digital native（数字原住民）的造词	13:44
chicken coop	/ˈtʃɪkən kuːp/	n.	鸡舍，鸡笼	18:04
city-states	/ˈsɪti steɪts/	n.	城邦（古希腊雅典、中世纪佛罗伦萨等自治城市国家）	6:17
come across as	—	phr. v.	给人以……的印象，显得（come across as too critical 显得过于苛刻）	28:25
commit murder	—	phr.	杀人，犯谋杀罪（commit + 罪行名词的固定搭配）	8:31
consciousness	/ˈkɔːnʃəsnəs/	n.	意识（哲学与认知科学核心概念）	41:26
costly	/ˈkɒːstli/	adj.	代价高昂的（不仅指金钱，也指时间精力）	30:32
deep fakes	/ˈdiːp feɪks/	n.	深度伪造：用AI合成的以假乱真的音视频、图像	29:56
dictate	/ˈdɪkteɪt/	n.	命令，指令。注意名词重音在前，动词 /dɪkˈteɪt/ 重音在后	5:25

embedded	/ɪm'bedɪd/	adj.	嵌入的，内嵌于（be embedded within/in）	28:25
emotive	/ɪ'moʊtɪv/	adj.	动情的，引发情感的	43:59
empathy	/'empəθi/	n.	同理心，共情	43:59
entities	/'entətɪz/	n.	实体（法律、抽象意义上独立存在的事物）	21:39
establishment	/ɪ'stæblɪʃmənt/	n.	建制，体制（the scientific establishment 科学界建制）	31:53
evoked	/ɪ'vəʊkt/	v.	唤起，引起（情感、记忆）（evoke 的过去分词）	40:48
exploit	/ɪk'splɔɪt/	v.	剥削；开发利用（此处指殖民剥削）	22:17
fact-checking	/'fækt tʃekɪŋ/	n.	事实核查（新闻业术语）	30:32
fantasize	/'fæntəsaɪz/	v.	幻想，想入非非	33:12
far-right	/,fɑːr 'raɪt/	adj.	极右翼的	20:48
feed	/fi:d/	n.	（社交媒体的）信息流，动态流	21:39
fertile	/'fɜːrtl/	adj.	肥沃的，富饶的	25:00
financiers	/,fɪnən'saɪz/	n.	金融家，投资人	37:18
flattering	/'flætərɪŋ/	adj.	奉承的，讨好的；使人显得更好的	31:13
foretaste	/'fɔːrteɪst/	n.	预先的体验，先兆（a foretaste of sth. 预尝……的滋味）	27:43
forge	/fɔːrdʒ/	v.	锻造；缔造（关系）；伪造。此处 forge intimacy 指制造亲密感	0:49
generous estimate	—	phr.	宽松的估计、往多了算（反义：conservative estimate 保守估计）	36:26
genie	/'dʒiːni/	n.	（神灯里的）精灵；源自阿拉伯民间故事	45:23
get stuck into	—	phr.	开始认真投入做（某事）；英式口语，常指开始正题或大快朵颐	0:02
glorified	/'glɔːrɪfaɪd/	adj.	被美化的；a glorified X = 说穿了不过是个X（贬义）	39:26
held power in check	—	phr.	制衡权力（hold sth. in check 抑制、约束）	1:34
hunter-gatherer	/,hʌntər 'gæðərər/	n.	狩猎采集者（农业出现前的人类生存方式）	33:58
immense	/ɪ'mens/	adj.	巨大的，极大的	20:48
imperial hub	—	phr.	帝国中枢（imperial /ɪm'piəriəl/ 帝国的；hub 枢纽）	24:02
impersonate	/ɪm'pɜːrsəneɪt/	v.	冒充，假扮（某人）	18:56
inherent contradiction	—	phr.	内在矛盾（inherent /ɪn'hɪərənt/ 固有的、与生俱来的）	9:37

intimacy	/'ɪntɪməsi/	n.	亲密，亲密感	0:49
intimate	/'ɪntɪmət/	adj.	亲密的 (intimate relationships 亲密关系)	41:26
kill switch	—	n.	紧急停止开关；可远程使系统彻底失效的「一键切断」装置	25:55
ledgers	/'ledʒərz/	n.	账本，分类账（会计术语）	3:11
legal fiction	—	n.	法律拟制：法律上「假定为真」的虚构安排，如公司人格	16:38
legal personhood	—	n.	法律人格（法律上作为权利义务主体的资格）	14:21
legionaries	/'li:dʒənəriz/	n.	（古罗马）军团士兵	26:46
lulled	/lʌld/	v.	哄骗使放松警惕 (lull sb. into sth. 使人不知不觉陷入)	43:59
material substrate	—	phr.	物质基底、物质载体 (substrate /'sʌbstreɪt/ 哲学与生物学术语)	43:10
medieval	/,medi'i:vəl/	adj.	中世纪的	6:17
metropolis	/mə'trɑ:pəlɪs/	n.	大都会；（相对殖民地而言的）宗主国中心	24:02
misery	/'mɪzəri/	n.	痛苦，悲惨 (the sources of misery 痛苦的根源)	32:39
mortal	/'mɔ:rtl/	adj.	终有一死的（反义：immortal 不朽的）	38:39
muse	/mju:z/	v.	沉思；若有所思地说 (muse that...)	37:18
no match	—	phr.	根本不是对手 (be no match for sb. 完全敌不过)	23:14
olive groves	/'ɑ:lɪv grəʊvz/	n.	橄榄园 (grove 小树林、果园)	24:02
persuasion	/pə'sweɪʒən/	n.	说服（力）(powers of persuasion 说服能力)	37:18
pharmaceuticals	/,fɑ:rmə'su:tɪkəlz/	n.	制药业；药品	46:38
poured into	—	phr.	（资金等）大量投入 (pour money into 砸钱于)	9:05
precedent	/'presɪdənt/	n.	先例，前例（法律与历史语境常用）	1:34
prompt	/pra:ɪmpt/	v.	促使，引发 (feelings that prompt the words 促使话语说出的感受)	40:48
reduced to	—	phr.	被贬低/沦落到.....的境地 (be reduced to the position of...)	35:40
rise up to	—	phr. v.	迎接、胜任（挑战）（更常见形式为 rise to the challenge）	46:01
rubber plantations	/'rʌbə plæn'teɪʃənz/	n.	橡胶种植园 (plantation 常指殖民地大种植园)	25:00
scramble	/'skræmbəl/	n.	争夺，抢夺 (a scramble for/to power 权力争夺战)	0:02

scriptures	/ˈskɹɪptʃərz/	n.	宗教经文，经典	3:11
shareholders	/ˈʃerhoʊldərz/	n.	股东	12:07
side dish	—	n.	配菜；喻指附属品、次要之物	7:39
sophisticated	/səˈfɪstɪkətɪd/	adj.	复杂精妙的，高度发达的	3:11
sovereignty	/ˈsɑːvrənti/	n.	主权；自主权	28:25
spears	/spɪrz/	n.	长矛，标枪	23:14
speculation	/ˌspekjəˈleɪʃən/	n.	推测，猜想（无实证的理论）	3:11
sphere	/sfɪr/	n.	领域，范围（the financial sphere 金融领域）	12:07
subservient	/səbˈsɜːrvɪənt/	adj.	卑躬屈膝的，唯命是从的	29:09
sue	/suː/	v.	起诉，控告（sue sb. in court）	15:52
syllogism	/ˈsɪlədʒɪzəm/	n.	三段论（由大前提、小前提推出结论的逻辑推理形式）	38:39
take over	—	phr. v.	接管，接手控制	13:44
the fox enters the chicken coop	—	phr.	狐狸进了鸡舍；喻指危险者获准进入其可肆意破坏之地	18:04
trajectory	/trəˈdʒektəri/	n.	轨迹，发展轨道（the trajectory of history 历史的轨迹）	33:12
tribesmen	/ˈtraɪbmən/	n.	部落成员，部落民	26:46
trustees	/trʌˈstiːz/	n.	受托人，托管人（法律术语）	12:07
unprecedented	/ʌnˈpresɪdəntɪd/	adj.	史无前例的，空前的	12:07
vassal	/ˈvæsəl/	n.	封臣；附庸（国）	29:09
yearning	/ˈjɜːrniŋ/	n.	渴望，向往（yearning to do / for sth.）	33:12

第四部分 理解自测 · 8 题

1. 如果能穿越回历史上任何时刻，赫拉利选择哪个时期？他给出的理由是什么？
2. 赫拉利如何论证「现代之前不存在大规模民主」？信息技术在其中扮演什么角色？
3. 工具与能动者(agent)的区别是什么？赫拉利用了哪两个历史例子说明？
4. 文中提到哪个国家宣布将赋予AI法律人格？赫拉利为什么认为这个决定「一旦做出就无法收回」？
5. 赫拉利指出AI帝国与罗马帝国、大英帝国相比有哪两个史无前例的特征？他举了什么现实案例？
6. 真相在与虚构的竞争中有哪三重劣势？人类历史上如何应对？
7. 赫拉利如何反驳「AI只是高级自动补全」的说法？
8. 赫拉利希望后人把这个时代称为什么？他引用了哪个关键数据说明当前AI发展的失衡？

参考答案

1. 石器时代的认知革命时期（约5-7万年前）。因为那是人类历史上最迷人也最不被理解的时刻：人类在极短的演化时间里凭借语言从不起眼的动物变成地球主宰。（第3-4段）
2. 民主本质是对话，古代只有雅典、佛罗伦萨等小规模民主，因为数百万人无法实时对话；直到印刷、报纸、广播、电视、互联网出现，大规模民主才成为可能——信息技术是民主的地基而非配菜。（第8-10段）
3. 工具不能自主决策、不能发明新事物；能动者两者都能。例子：印刷机不能自己决定印《圣经》还是《古兰经》（由古腾堡决定），原子弹不能自己选择轰炸广岛还是东京、也不能发明氢弹；而AI武器既能选目标又能发明下一代武器。（第14-15段）
4. 阿根廷。因为AI是「官僚体制的原住民」，一旦获得开户、投资、诉讼等法律权力，就能凭借远超人类的能力接管金融、法律等控制系统——「狐狸进了鸡舍」，人类无力夺回。（第20、23-24段）
5. 一是权力可完全集中：信息和代码可全部集中于一两个国家，而土地、油田、种植园搬不走；二是存在kill switch：帝国中枢可远程让卖出的技术失效，而卖出的钢剑和步枪无法遥控。案例是星链——马斯克按下按钮，乌克兰无人机就失灵。（第31-35段）
6. 真相昂贵（需要时间精力核查）、复杂（人们喜欢简单故事）、痛苦（人们不想知道关于自己的某些事实）；虚构则廉价、简单、讨好。人类的应对是特意建立求真机构，如新闻业和科学建制。（第39-41段）
7. 他通过内省指出，人脑说话时似乎也在「预测下一个词」——他开口说一句话时并不知道它如何结尾，也不知道词从何而来。如果思考就是排列词语，那AI与人并无本质区别；真正的分歧在于思考是否需要感受，而AI能否感受，科学尚无法检验。（第51-54段）
8. 「智慧革命」——智能（实现愿望的精灵）变得廉价后，瓶颈变成选择愿望的智慧。数据：AI公司在安全与提升能力上的投入约为1:100，每花1亿美元让AI更强大，只花100万确保安全，其他行业都不会容忍这种比例。（第58-61段）