Promoting Global Solidarity Through AI

（《同球共济与人工智能》）

**Hello, friends,**

**观众朋友们大家好**

**welcome to TechWorld.**

**欢迎来到《科技π》**

**I am your host, Wu Baojun**

**我是主持人吴宝俊**

**from the University of Chinese Academy of Sciences.**

**来自中国科学院大学**

**Today, our program is on**

**我们今天的话题叫做**

**"Promoting Global Solidarity Through AI".**

**同球共济与人工智能**

**We have two experts with us**

**请来了两位专家**

**to share their insights.**

**与大家探讨人工智能领域的话题**

**Sitting**

**我们第一位专家**

**opposite me**

**坐在我对面的**

**is Professor Zeng Yi, a researcher from the Institute**

**of Automation, Chinese Academy of Sciences.**

**来自中国科学院自动化研究所的研究员曾毅**

**Professor Zeng is also**

**同时曾毅老师也是**

**an expert of the United Nations' high-level advisory body on AI.**

**联合国人工智能高层顾问机构专家**

**Hello, everyone.**

**大家好**

**The second expert,**

**第二位**

**sitting next to me,**

**坐在我旁边的**

**is Professor Gao Shaolin, an expert at the Peking University Law**

**and AI Research Center.**

**是北京大学法律与人工智能研究中心专家顾问高绍林老师**

**Our first topic**

**那么首先我们进入第一个话题**

**is a popular science topic.**

**是一个科普类型的话题**

**We know that**

**我们知道**

**China has made great strides in developing AI.**

**中国在人工智能领域有了很多先进的技术和应用的案例**

**We would like you to**

**我们希望二位专家**

**share some examples with our audience.**

**先给我们屏幕前的公众介绍一些这方面的例子**

**So why should we develop AI technology?**

**我们为什么要去发展人工智能这样的技术**

**I think that is quite crucial.**

**我觉得这是非常关键的**

**AI applications free us**

**就是人工智能技术当然希望我们人类**

**from repetitive labor.**

**从这种纷繁重复性的劳动当中解脱出来**

**For instance, if we develop AI**

**比如说我们未来发展人工智能**

**and robotics technology,**

**机器人技术**

**we can use them in hazardous fields,**

**那么在一些这个危险的领域**

**like firefighting,**

**比如说救火 消防**

**or even**

**甚至比如说**

**when exploring outer space,**

**在我们去探索外空的时候**

**where we can't send many scientists**

**可能没有办法说去那么多的科学家**

**or explorers.**

**或者是我们的探索者**

**At such stages,**

**这时候你可以**

**the combination of AI and robotics helps,**

**人工智能与机器人技术的结合**

**such as the development of our lunar rovers.**

**比如我们的月球车的发展**

**These rovers have AI recognition**

**现在上面就是有很多人工智能的识别**

**and multi-sensory capabilities,**

**多感觉的这种理解等等的这种能力**

**enabling us to accomplish tasks we might struggle with.**

**使得我们很多人不善于做的事情都可以用人工智能去完成**

**Professor Gao,**

**高老师**

**in our daily life,**

**在咱们日常生活当中**

**what are some impactful examples of AI application?**

**您觉得有哪些是您感受最深刻的**

**Recently, I took a 3,000-kilometer trip on a highway.**

**最近我做了一趟旅行 开了3000公里的高速公路**

**In the past, while driving such a long distance,**

**以往3000公里的高速公路**

**it was common to come across traffic accidents.**

**大概看个三五起交通事故应当很平常**

**But on this trip,**

**但是我这3000公里跑下来之后**

**I didn't see any accident at all.**

**居然一个事故也没有看到**

**I think the reasons are two-fold.**

**其中我就想原因可能有两个**

**First, people have become more aware of safety.**

**第一 就是人们的安全意识提高了**

**Second, many cars now have assisted driving systems.**

**第二 就是我们现在许多汽车都有了辅助驾驶系统**

**Even trucks, like the ones from FAW,**

**连我们解放牌大卡车 一汽生产的卡车**

**have this technology.**

**大卡车上它也有辅助驾驶技术**

**If the driver grows tired, the system will alert them.**

**司机疲劳它会提醒**

**So in our daily lives,**

**所以我们在生活当中就会发现**

**we have seen a significant reduction in accidents.**

**安全事故大大减少了**

**And I think AI plays a huge role in this.**

**其中我觉得人工智能技术功不可没**

**Professor Gao mentioned driving.**

**刚才高老师提到一个开车上路的这么一个话题**

**I recently visited Yizhuang in Beijing,**

**我刚好前一段去了一趟北京亦庄**

**which is an autonomous driving test area.**

**那边好像是一个自动驾驶的试验区**

**I saw cars in the streets with no drivers,**

**看这个大街上跑着的这个车呀**

**even garbage trucks.**

**过去一个垃圾车没有人开**

**These vehicles are all autonomous.**

**还有去的各种各样车都是没有人开的**

**咱们国家在自动驾驶技术领域的发展**

**So what is the current status**

**of China's autonomous driving technology?**

**目前是怎么样的一个状态**

**It's still in the testing phase, not yet fully rolled out.**

**在测试阶段 还没有正式地全面推开**

**这个本身跟我们原来对汽车的概念**

**This is very different from our original concept of cars.**

**是有一个很大的差异的**

**Previously, we thought of vehicles**

**原来我们对汽车**

**and non-motorized transport**

**对于非机动车**

**as human-driven.**

**我们都是以人为主的**

**We determined responsibility based on the driver's actions,**

**我们判别的责任是看驾驶员的责任**

**the cyclist's behaviors**

**看我们骑自行车**

**and the actions of pedestrians.**

**我们行人的责任**

**In the future,**

**那么未来**

**with autonomous vehicles,**

**自动驾驶汽车出来之后**

**there will be no human driver,**

**它确实是没有人了**

**none at all.**

**没有人**

**AI is actually a tool created by humans.**

**实际是人做的一个工具**

**It's essentially an automated tool running by itself.**

**它相当于一个自动工具在运行**

**Currently, our smart connected vehicles**

**所以咱们现在智能网联汽车**

**are in the testing phase.**

**叫测试阶段**

**It is not just a technical test,**

**它测试不仅仅是技术上的测试**

**but also a test of the allocation of future rights and responsibilities.**

**其实还有未来的**

**权利义务关系的分配上的这种测试**

**How can technology best serve humanity?**

**怎么样去 使科技更好地为人服务**

**By humanity,**

**这里的人**

**I don't mean individuals,**

**当然不是一个个体的人**

**but those who liberate productivity,**

**是整个解放生产力的同时**

**and contribute to the development of autonomous driving.**

**都为无人驾驶作出贡献的这些人**

**How can they receive corresponding benefits**

**他们怎么能够从科技发展当中**

**from technological advancements?**

**也能够得到相应的回馈**

**This is an issue that I think should be considered**

**in the testing process.**

**这是一个我觉得测试当中也应当考虑的问题**

**However, right now, we focus only on**

**但是我们现在往往测试**

**testing and improving the technology.**

**想的只是技术上的测试和完善**

**In fact, many legal aspects will also need to be refined in the future.**

**实际未来有很多法律关系也是要完善的**

**I think**

**我印象当中**

**one of the early explorers of AI**

**人工智能早期的探索者之一**

**was Edmund Berkeley.**

**1949年的时候埃德蒙·伯克利**

**In 1949, he wrote a book called**

**他写过一本书叫做**

**"Giant Brains, or, Machines That Think".**

**《巨脑：可以思考的机器》**

**In the last chapter,**

**它的最后一章**

**he talks about what (happens) if a machine is truly intelligent**

**他讲如果机器变得真正智能了**

**(and) the social control problems.**

**会出现什么社会问题**

**就是讲如果我们真的构造了一个**

**It talks about how to control**

**an intelligent machine if we really build one.**

**智慧的机器 如何去对它进行控制**

**There is a passage in it**

**它那里面有一段话**

**that I really appreciate.**

**我非常的欣赏**

**It says that for a scientist,**

**他说作为一个科学家来讲**

**the easiest thing is to make scientific discoveries**

**最容易的事情是做科学发现**

**or inventions.**

**或者是科学发明**

**The harder thing is to think**

**比较难的事情是想**

**about the potential negative impacts of these discoveries**

**or inventions on society.**

**你的这种发现或发明对于社会可能造成的潜在的负面的影响是什么**

**So I think,**

**所以我觉得**

**for innovators of modern AI technology,**

**这个对于现代的这种人工智能的科技创新者**

**this puts forward higher requirements.**

**其实也提出了更高的要求**

**Both of you are responsible scholars.**

**我看得出二位老师都是负责任的学者**

**Professor Zeng Yi**

**曾毅老师作为**

**is a researcher at the Institute of Automation,**

**Chinese Academy of Sciences,**

**中国科学院自动化研究所的研究员**

**an expert in the field of AI.**

**一个人工智能领域的专家**

**When discussing AI,**

**在讲人工智能话题的时候**

**you don't just talk about its broad prospects,**

**不是一味地说人工智能有非常广阔的前景**

**or just focus on its advantages.**

**不是单纯地讲它的这个优点**

**You also remind the public**

**同时也要提醒大家**

**that AI is not perfect**

**人工智能并不完美**

**We should always advocate**

**我们一直要倡导 负责任地发展人工智能技术**

**responsible development of AI technology.**

**负责任地发展AI**

**How do we understand this responsibility?**

**那么如何理解这个负责任**

**I remember that in 2019,**

**我记得在2019年的时候**

**the National Governance Committee for the New Generation AI**

**国家新一代人工智能治理委员会**

**released the Governance Principles for**

**New Generation AI —Developing Responsible AI**

**当时发布新一代人工智能治理原则**

**The subtitle was**

**副标题就叫做**

**"Developing Responsible AI".**

**发展负责任的人工智能**

**Many scholars from different walks of life**

**有很多来自于社会的不同界的学者**

**had discussions with us.**

**就跟我们讨论**

**I remember there was a philosopher who asked me,**

**我记得当时有一个哲学家讨论**

**Professor Zeng,**

**说曾老师**

**when you wrote about developing responsible AI,**

**你们写的这个发展负责任的人工智能**

**were you asking AI to be responsible?**

**你是让人工智能负责任吗**

**I said of course not.**

**我说当然不是这样**

**I said it has both short-term**

**我说它有一个近期的意义**

**and long-term implications.**

**有一个长远的这个意义**

**First,**

**首先第一个**

**AI right now**

**现在的人工智能**

**is just a tool for seemingly intelligent information processing.**

**它只是一个看似智能的信息处理的一个工具**

**So when we talk about responsible AI,**

**所以现在我们说负责任的人工智能**

**we must be very clear**

**我们要非常清晰**

**that the responsibility lies with**

**这个责任在我们的**

**the developers and operators of the product and service.**

**研发产品和运营服务的部分**

**The responsibility is to be borne by the persons deploying it.**

**这个责任是要部署者来承担的**

**This issue is also being discussed in the legal community.**

**这个问题其实是在法律界也在讨论**

**It is about the entity of AI,**

**就是关于人工智能的主体性问题**

**whether it can be considered a social entity.**

**到底它是不是一个社会主体**

**This is what I am curious about too.**

**这也是我很好奇的**

**Legally, should AI be treated as an entity that**

**can bear legal responsibility?**

**在法律上 它是不是应该可以当做一个能负法律责任的主体**

**Since 2019,**

**从我们2019年**

**when China released the Governance Principles**

**for New Generation AI — Developing Responsible AI,**

**咱们发布负责任人工智能指导意见以后**

**there has been much discussion.**

**其实这个讨论还是很多的**

**Personally,**

**我个人来说**

**I believe AI cannot be a legal entity**

**in the foreseeable future.**

**我觉得人工智能在我们现在的可预见的未来**

**还不能成为法律上的主体**

**As Professor Zeng said,**

**刚才曾老师说的**

**the responsibility lies with the developers and operators,**

**就是要由开发者 运营者来负责任**

**not with AI itself.**

**而不是由人工智能体本身来负责任**

**If AI were a legal entity,**

**因为人工智能体如果作为一个法律上的主体存在**

**it would be fundamentally different from the legal**

**entities we currently have.**

**那么它的主体性 跟我们现在所讲的法律上的主体 有本质上的差异**

**If it committed a crime,**

**它犯罪了**

**you can't sentence it to death.**

**你把它判了个死刑**

**But it would still be a mere tool.**

**它就是一个工具**

**In essence, it is just a tool.**

**它本质上还是工具**

**So I think despite the debate on this issue,**

**所以我觉得这主体性问题尽管可以讨论**

**the mainstream view**

**但是我觉得现在主流的观点**

**among most scientists,**

**和绝大多数的科学家**

**legal scholars, and sociologists**

**我们法学家和社会学家 可能还都是认为它就是一个工具**

**is that AI is just a tool, not an entity.**

**它不具有主体性**

**After hearing your views,**

**听了二位老师的观点**

**I am reminded of a saying.**

**我忽然间想起一句话**

**There used to be a famous saying on the Chinese Internet:**

**过去在中文的互联网上有一句名言**

**"Don't take the initiative, don't refuse,**

**and don't take responsibility."**

**叫做不主动 不拒绝 不负责**

**But in the context of robots**

**但是套在机器人**

**and AI,**

**套在人工智能领域**

**it should be very proactive,**

**neither refusing nor taking responsibility.**

**它应该是很主动 不拒绝 也不负责**

**Why no responsibility?**

**为什么不负责**

**Because AI is essentially a tool to assist humans.**

**因为它本质上还是一个人的辅助的工具**

**It cannot replace humans.**

**不能够取代人**

**So from this perspective,**

**所以从这个角度讲**

**although AI appears very intelligent on the surface,**

**人工智能尽管前台很智能**

**the real control still lies with humans.**

**但后台真正控制的还是应该是人工**

**The more intelligent it is,**

**the more the human control over it.**

**它有多智能就应该有多人工**

**Do you agree with this?**

**二位老师 看我这说的对不对**

**Yes, I think the Chinese translation of the term AI**

**对 我觉得我们中文翻译AI这个词**

**is very accurate and clever.**

**用的这四个字就很准确和精巧**

**Its premise is that AI is artificial and human-empowered.**

**它的前提是人工**

**Humans make the machine intelligent.**

**是人工使这个机器产生的智能**

**So, who is responsible?**

**那么谁来负责任**

**Humans are responsible,**

**是人来负责任**

**not the machines.**

**而不是让机器去负责任**

**We must understand that**

**我们要知道**

**humans are not the most intelligent beings.**

**人类并不是意识水平最高的智能体**

**And if we reach that stage,**

**如果达到了这样阶段的话**

**then the issues we just discussed**

**那我们刚才讨论的这些问题**

**need to be reconsidered.**

**都需要重新的思考了**

**For example, I asked you a question.**

**比如我问宝俊一个问题**

**When superintelligence really arrives,**

**就是你觉得超级智能真正到来的时候**

**will it be willing to sit down**

**它愿不愿意坐着**

**and have discussions with you calmly?**

**跟你心平气和地来讨论问题**

**Or, through human wisdom,**

**或者说你通过人类的这个智慧**

**can you balance**

**你能不能够制衡**

**or harmoniously coexist with an intelligent being**

**或者说是跟一个智能水平**

**that far surpasses your own intelligence?**

**远远超越你的智能体和谐共生**

**On this issue,**

**就这个问题本身**

**can humans have an answer in 10 or 20 years?**

**人类能不能够在10年20年之内准备好答案去应对**

**This is the issue we need to focus on.**

**这个是我们需要去关心的问题**

**In my personal view,**

**在我个人看来**

**I could never accept an intelligent being**

**我是绝无可能接受有一个**

**that surpasses myself, surpasses human intelligence,**

**超越自己 超越人类的智能体**

**controlling our lives.**

**去掌控我们的生活的**

**If there is an intelligence that surpasses us,**

**如果有一种智能超越了我们人本身的话**

**then, not to mention the potential dangers and harms,**

**那么对我们的生活且不说潜在的危险和伤害**

**happiness would be lost, to say the least.**

**至少快乐就没有了**

**I remember vividly the time in 2016,**

**我印象当中非常深刻的是2016年的时候**

**when Google's AlphaGo had just beaten Lee Sedol.**

**当时谷歌的阿尔法狗刚刚战胜了这个李世石**

**I attended a conference at the University of Leeds in the UK,**

**后来我在这个英国利兹大学开会**

**and I saw several elderly men**

**在马路上看到好几个老头在那**

**playing chess in the street.**

**下国际象棋**

**I went over to them and asked a question**

**我过去我半挑衅性地问了一个问题**

**I wanted to actually hear their thoughts,**

**我的那种挑衅实际上是希望听到他们的一些反思**

**not provoke them.**

**不是真的去挑衅他们**

**I said, "Do you know about the latest developments in AI?**

**我说你们知不知道人工智能的最新进展**

**In 1997,**

**1997年的时候**

**IBM's Deep Blue**

**这个IBM深蓝**

**beat Garry Kasparov.**

**就已经下过了卡斯帕罗夫了**

**And this year, Google's AlphaGo**

**然后在今年的时候 谷歌的阿尔法狗**

**has surpassed**

**已经下过了在围棋领域已经突破了**

**one of the highest human levels in Go."**

**人类的这个当时最高水平之一了**

**I said, "You gentlemen playing chess here,**

**我说你们在这下这个国际象棋是有什么**

**what are your thoughts on this?"**

**你们的考虑是什么**

**Those old gentlemen looked at me and smiled.**

**那几个老先生就笑着就看我**

**他们就说他说这个年轻人**

**One of them said,**

**"Young man, what has that got to do with us?"**

**他说这东西跟我们有什么关系**

**Actually,**

**实际上就是我想说的意思**

**what I wanted to say was that**

**就是这样的**

**AI may surpass human intelligence.**

**就是人工智能的智慧超越人类**

**So the key question is**

**但是最关键的是**

**will AI be able to coexist harmoniously with humans in future,**

**人工智能未来是不是能与人类和谐共生**

**rather than have a competitive relationship?**

**而不是一个竞争的这个关系**

**Chess, Go, and video games,**

**国际象棋 围棋 电子游戏**

**if AI has surpassed humans in these areas,**

**它下过了**

**that's its problem.**

**它下过是它的问题**

**But your life, your happiness,**

**你的生活 你的愉悦**

**they come from our own inner needs.**

**那是发自我自身的这种需求**

**Even now, people are exploring the possibility of**

**就哪怕说现在大家都在探索说**

**the emergence of AI scientists in the future.**

**未来是不是有人工智能科学家的这种可能**

**For me, this has no impact.**

**我觉得这个对我都没有任何的影响**

**I choose to be a scientist**

**因为我选择做科学工作者**

**because of the joy I gain**

**那是因为我进行科学突破**

**when I make scientific breakthroughs.**

**我自己产生的愉悦感**

**It's not about the achievements I've made**

**而不是我将取得的什么样的成就**

**or what I've exchanged for them.**

**换什么东西**

**Actually, I think**

**实际上我觉得**

**human thought,**

**人类的思想**

**and how humans respond to**

**我觉得人类如何应对**

**the ecology and future challenges**

**生态和未来的一些态度 我觉得也需要随着**

**need to evolve and accelerate with time.**

**演化的步伐加速演进**

**Professor Gao,**

**高老师**

**what do you think?**

**您对这个问题怎么看**

**At this stage,**

**那么在这个阶段**

**when we give intelligence to machines,**

**我们赋予这个机器智能的时候**

**maybe we need to also**

**incorporate basic human principles and ethics,**

**可能就要把人的这些基本理念 基本道德**

**the most fundamental ethics of our society,**

**就是我们伦理当中最基本的东西**

**and make AI accept them,**

**能够使人工智能接受**

**and strictly follow these principles and ethics as it develops.**

**并且在发展中 它也能够严格地遵循着这套伦理向前发展**

**This is to ensure that**

**这就是说**

**in the future, the intelligence of machines**

**未来的机器的智能**

**doesn't harm our planet.**

**它不至于危害我们这个星球**

**That is not the result human technological**

**advancements want to achieve.**

**这个结果不是我们人类科技所要达到的结果**

**So, going back to today's topic.**

**所以我们回到今天说**

**While we develop AI,**

**那么我们在研发人工智能**

**and focus on technology,**

**在关注技术的同时**

**we must also focus on its fundamental ethics.**

**必须关注它的基础伦理**

**And these fundamental ethics**

**而且这些基础伦理**

**must be embedded into its intelligence,**

**要嵌入它那个智能里面去**

**into its DNA,**

**在它“DNA”里边**

**with built-in self-restraint.**

**就带有这种自我约束**

**We must develop responsible AI,**

**我们发展负责任的人工智能**

**which means throughout the process of AI development,**

**就是在人工智能发展过程当中**

**we must uphold the concept of a community with**

**a shared future for mankind,**

**要自始至终用人类的 叫做我们人类命运共同体的**

**and our shared human values to train it.**

**这个理念 用我们人类共同价值观来训练它**

**Then it can develop intelligence**

**使它产生智能**

**and coexist harmoniously with humanity and other species,**

**最后达到它与人类 与其他物种**

**都能够和谐共生**

**ultimately maintaining the balance of our planet's ecosystem.**

**保持我们这个星球 整个生态的平衡**

**This may be our ideal.**

**这可能是我们人的理想**

**That's why I say**

**所以我说**

**China's proposal of responsible AI**

**我们中国较早提出负责任的人工智能**

**is a very visionary concept.**

**我觉得是非常有远见的一个理念**

**In the course of the conversations,**

**在对话的过程当中**

**naturally,**

**自然而然的**

**everyone started discussing the safety,**

**大家就探讨起了人工智能的安全性**

**reliability,**

**可靠性**

**and the ethical framework issues of AI.**

**以及它在AI的伦理框架当中的一些问题**

**Professor Gao also mentioned**

**那么高老师也提出来了**

**that humans should coexist harmoniously with AI.**

**说人要和人工智能要和谐地共存**

**And Professor Zeng's perspective,**

**而曾老师这个观点**

**in my opinion, provides me with a new angle.**

**在我看来也更新鲜**

**In fact, he mentioned that**

**其实他提到了一点**

**就是人对于未知的事物**

**people often feel anxiety and fear about unknown things.**

**就是人对于未知的事物 往往会有一种焦虑和一种恐惧感**

**This fear**

**这种恐惧感**

**makes us defensive,**

**让我们有一种防卫意识**

**which comes from the insignificance of humans.**

**这源自于我们人的渺小**

**So, from this perspective,**

**所以从这个角度讲**

**we are researching AI**

**我们研究人工智能领域**

**to explore unknown areas,**

**其实就是为了把那些未知的地带**

**to unleash its full potential,**

**给它探索开**

**to eliminate human fear,**

**把这个事看透了 以消除人的恐惧感**

**and to make AI more reliable.**

**让人工智能更为可靠**

**So, regarding the development of AI,**

**咱们国家在人工智能领域这个发展**

**what is China's domestic policy?**

**国家对内的政策是怎么样的**

**And what is our basic attitude**

**以及我们在对外交流合作过程当中**

**towards international cooperation in this field?**

**我们提出的 对外的合作的基本态度是什么**

**Let's talk about this.**

**咱们接下来聊聊这个话题**

**In China,**

**我们国家的**

**the AI development strategy**

**人工智能的发展**

**is consistent and closely integrated both**

**domestically and internationally.**

**应该说它的发展战略 是对内和对外结合的**

**In fact, our country has only one plan,**

**实际上我们国家只有一套**

**the Next Generation AI Development Plan,**

**叫做国家《新一代人工智能发展规划》**

**which was released in 2017.**

**这个是在2017年发布的**

**I think it has many highlights.**

**其中我觉得还是有很多的亮点**

**For example, it doesn't follow**

**比如说并没有停留在**

**the same approach as developing AI,**

**按照同一个脉络去发展人工智能**

**such as using big data,**

**比如说你采用大数据的方式**

**or traditional AI methods**

**like knowledge representation and reasoning,**

**还是你采用知识表示推理的**

**a traditional AI development approach,**

**这种传统的人工智能方式**

**or using modern deep learning**

**还是采用这个现代的深度学习**

**and large models.**

**然后这个大模型的这种方式**

**It actually explores many different possibilities.**

**其实里面还是探索了很多不同的可能性**

**But regarding AI development**

**但是对于人工智能的发展**

**and its potential social issues,**

**它可能潜在的社会问题**

**it clearly addresses them.**

**其实也有了非常明确的阐述**

**There are at least a dozen places**

**至少有十几处**

**where social ethics are discussed,**

**谈到了社会伦理的问题非常重要**

**and it also mentions that**

**那么也谈到了 就是未来**

**我们国家在法律方面会进行 人工智能相关的法治的建设**

**China will develop laws related to AI.**

**But when it comes to the vision part,**

**但是谈到愿景的部分**

**actually, in 2019,**

**其实这个是在2019年的时候**

**the Governance Principles for the New Generation AI —**

**Developing Responsible AI [was released].**

**国家《新一代人工智能治理原则》发布的时候**

**An international seminar was held in Puerto Rico**

**在波多黎各开了一个国际研讨会**

**on developing beneficial artificial general intelligence (AGI).**

**发展有益的通用人工智能**

**The most important question they asked,**

**问到我最重要的**

**which I believe we need to reconsider, was:**

**在我看来我们回来要需要思考的问题**

**Where is China's AI going?**

**他们讲说中国人工智能去向何方**

**In May,**

**5月份回来**

**we wrote the Beijing AI Principles.**

**我们就写了叫做《人工智能北京共识》**

**It was released in Beijing,**

**就通过在这个北京发布的**

**and it was called the Beijing AI Principles.**

**叫Beijing AI Principles**

**Within seven days,**

**发布了以后 7天之内**

**MIT Technology Review reported**

**当时《麻省理工科技评论》有一个报道**

**that the Beijing AI Principles**

**他们说看《人工智能北京共识》**

**clearly addressed some issues.**

**就说清楚了有些问题**

**It said that if the principles for AI development from China**

**他说如果说来自于中国的 这些人工智能发展的原则**

**don't talk about protecting privacy,**

**不去谈保护隐私**

**or respecting human rights,**

**不去谈尊重人类的这个权利**

**then not cooperating with China is understandable.**

**那么不和中国合作那是理所应当的**

**But since the Beijing AI Principles clearly states**

**但是如北京共识写的这么清楚**

**that privacy should be protected**

**保护隐私 保护人类的**

**and human rights should be protected,**

**这个基本的权利**

**then not to engage with China**

**那么不和中国合作**

**on this matter is self-defeating.**

**是我们自己的问题了**

**In the Governance Principles for New Generation AI,**

**在国家《新一代人工智能治理原则》当中**

**we wrote very clearly**

**当时我们就写得就非常清晰了**

**who China is developing AI for.**

**你中国发展人工智能是为谁**

**If you're saying that AI helps develop the Chinese economy,**

**你说发展中国经济**

**it is natural that AI will empower**

**the economy's development.**

**人工智能赋不赋能中国发展经济**

**那是理所当然的**

**But our goal is not just that.**

**但是我们的目标就不止于此**

**Our goal is that the development of AI**

**我们的目标就设在了**

**must contribute to achieving global sustainable development goals.**

**就是人工智能发展**

**AI development**

**一定要推进全球可持续发展目标的实现**

**must empower the realization of a community**

**with a shared future for mankind.**

**一定要赋能人类命运共同体的这个实现**

**So, you mentioned about the vision part.**

**所以我想你刚才讲的这个愿景的部分**

**The Governance Principles for the New Generation AI**

**就国家《新一代人工智能发展规划》**

**should set the direction for development**

**这份文件应该是在 国家《新一代人工智能发展规划》的基础之上**

**based on the New Generation AI Development Plan.**

**为发展的方向就定了一个调**

**Professor Zeng,**

**曾老师**

**I really learned a lot from your answer just now.**

**刚才这个回答我是真是学习到了**

**China is really visionary.**

**我们这个站位真的是足够高**

**It is looking at the issue from a global perspective,**

**站在一个全球人类的这么一个视角 来看待这个问题**

**not just our own country.**

**而不仅仅是局限于我们自己的这一个国家**

**Yes.**

**是的**

**The global sustainable development goals**

**全球可持续发展目标**

**have an international consensus.**

**这个是有国际共识的**

**So, we too have to make these goals understood by others**

**所以你要还要让别人看懂**

**and follow international consensus**

**还要遵循国际共识**

**and respond to international consensus.**

**还要响应国际共识**

**So its stance**

**所以它的立意**

**is indeed from the perspective of a global vision.**

**确实是站在这个全球的视野上的**

**My personal interpretation is**

**所以我个人看起来一个明确的解读**

**that the goal of our country is not a selfish one.**

**就是我们国家的这个目标**

**并不是一个自私的目标**

**No.**

**不是**

**China is genuinely responsible**

**我们是一个真正有负责任的**

**and hopes to benefit the entire world.**

**希望能够惠及全球的**

**This is our development strategy.**

**这么一种发展战略**

**Yes.**

**是的**

**所以为什么落脚点**

**So, why is the focus on the realization of a community**

**with a shared future for mankind?**

**在全人类命运共同体的实现**

**This is completely consistent**

**with our country's development.**

**这跟我们国家的这个发展 是完全一致的**

**So this year, China proposed a resolution,**

**所以今年中国**

**which was co-sponsored by over 140 countries.**

**咱们主提 有140多个国家附议的**

**It was the UN resolution on**

**"Enhancing International Cooperation on Capacity-building of AI",**

**联合国的人工智能的决议**

**which I found particularly impressive.**

**我觉得给我印象特别深的**

**There are two points.**

**有两条**

**One is eliminating the digital divide.**

**一条 就是消除数字鸿沟**

**The second is helping developing countries grow.**

**第二条 就要帮助发展中国家发展**

**你不能发展了人工智能之后**

**You can't have AI development where**

**the strong get stronger**

**你不能发展了人工智能之后 强者更强**

**and the weak get weaker.**

**弱者更弱**

**That's not our goal.**

**这个不是我们的目标**

**So the resolution we proposed**

**所以我们提出这决议案**

**gained support from most countries in the world.**

**也得到了全世界 绝大多数国家的一致赞成**

**Domestically,**

**就我们国内来说**

**in terms of domestic regulations,**

**国内的这个规范来说**

**over the years,**

**这些年**

**while developing,**

**我们始终在发展的同时**

**we have always paid attention to setting boundaries for AI.**

**注意给人工智能划边界**

**For example, after AI emerged,**

**比如人工智能出来之后**

**the first issues were related to the Internet.**

**最初出现的是网络问题**

**So the National People's Congress passed**

**two resolutions regarding Internet security.**

**所以全国人大通过了 关于互联网安全的两个决议**

**Later, we formulated the Cybersecurity Law**

**后来我们又制定了《中华人民共和国网络安全法》**

**from the perspective of security.**

**就是从安全角度**

**Then, when data issues became prominent,**

**然后数据的问题突出出来以后**

**we formulated the Data Security Law.**

**我们又制定了《中华人民共和国数据安全法》**

**Later, when deepfake technology appeared,**

**后来深度伪造出来之后**

**we promptly issued a management measure**

**for deep synthesis.**

**我们又及时发了一个**

**关于深度生成的《互联网信息服务深度合成管理规定》**

**And when large models emerged,**

**那么大模型出来之后**

**we issued an interim regulation**

**on the management of generative AI services.**

**我们又出台了一个《生成式人工智能服务管理暂行办法》**

**This measure still requires**

**这个办法还是**

**registering the generative AI service provider first.**

**你先要主体备案**

**What these service providers are doing**

**needs to be known to society.**

**主体你在做这件事要让社会知道**

**Moreover, the Cyberspace Administration of China**

**而且这些主体**

**has made these entities public on its website.**

**国家网信办都在网站上给你公开**

**This is called transparency.**

**这叫透明度嘛**

**Then the models you have developed must be filed for record.**

**然后你已经开发的模型要报备案**

**To file for record means**

**people know who is developing these models.**

**就是说这些模型大家知道是他在开发的**

**When you develop a system,**

**那你开发这个**

**you must take responsibility for it.**

**你就要对这个系统要负责**

**In this way, we define boundaries**

**我们采取划边界（的方式）**

**and leave space for development**

**然后给发展留下空间**

**to gradually move forward.**

**来逐步地向前发展**

**This way, AI truly moves**

**这样使人工智能真正**

**towards a positive direction**

**朝着一种向善的**

**that contributes to forming a community**

**with a shared future for mankind.**

**有利于形成人类命运共同体的**

**这么一个方向来发展**

**The development of AI in recent years**

**看得出 人工智能领域最近几年的发展**

**has been truly impressive.**

**真的是飞速**

**Our country must keep pace**

**with the times in terms of specific policies**

**咱们国家在这个具体的政策上**

**and domestic regulations,**

**内部的规则制定上也要与时俱进**

**update them in real time**

**实时地更新它**

**to achieve a dynamic change and balance.**

**实现一个动态的变化和平衡**

**When we make real-time adjustments,**

**我们在实时调整的时候**

**we also encounter situations**

**也是遇到一种情况**

**that require us to adjust legislation accordingly.**

**我们随之调整立法**

**This approach is called agile governance.**

**所以这种方式叫做敏捷治理 敏捷治理**

**What does agile governance mean?**

**敏捷治理是什么**

**It is a key term used in the principles of modern AI governance.**

**就是我们《新一代人工智能治理原则》的关键词**

**When you formulate a law,**

**一个法律那么你制定**

**you need to implement it.**

**你要去实施**

**But technology keeps evolving,**

**但是技术在不断地演进**

**so how do we carry out this kind**

**所以怎么去进行这种**

**of adaptive governance?**

**自适应的治理**

**After creating a regulation, you can't say that**

**你不是说制定了这个东西**

**it will remain unchanged for 10 years.**

**你以后十年之内不变**

**That's impossible.**

**那是不可能的**

**The AI field is booming,**

**就是说人工智能领域在蓬勃发展**

**but its development is not a wild and uncontrolled expansion.**

**但是它的这个发展并不是野蛮生长**

**It's not like weeds spreading unchecked.**

**它不是杂草**

**We hope it grows like a well-tended crop.**

**而是我们希望它像一个庄稼一样 茁壮地成长**

**So, we need to guide it appropriately**

**所以我们需要合适地去引导它**

**and regulate it to serve the interests of all humanity.**

**把它规范到服务我们全人类的利益**

**This is the path we should take.**

**这个路线上来**

**AI governance,**

**人工智能的治理**

**including ethical norms and rule of law,**

**包括伦理规范 法治**

**will only help AI develop steadily.**

**它只会帮助人工智能稳健地发展**

**What it actually does**

**但是它做的事情**

**is to prevent the uncontrolled growth of AI.**

**确实是防止人工智能野蛮地生长**

**The world doesn't need**

**这个世界上不需要**

**every industry, every place,**

**所有的行业 所有的地方**

**to have AI everywhere.**

**都让人工智能无处不在**

**It's not necessary.**

**不需要的**

**In many fields,**

**就很多地方**

**you don't need AI at all.**

**你根本不需要人工智能**

**Leave those fields to humans.**

**不需要人工智能留给人**

**I'm curious about**

**那么我比较好奇的是**

**other countries' basic attitudes in this field.**

**其他国家在这个领域当中 他们的基本态度**

**Do they have**

**他们有没有一些提出来的**

**an approach similar to ours?**

**类似于我们这样的方针**

**What similarities**

**和我们的有哪些相同点**

**and differences do they have compared to us?**

**和不一样的地方**

**Could you elaborate on this?**

**二位专家可以介绍一下吗**

**We collected**

**我们这个收集了**

**all the ethical principles of AI published worldwide,**

**全球发布的所有的人工智能的 这个伦理原则**

**and conducted a quantitative analysis.**

**然后进行了一个量化的分析**

**It shows that 95% of the topics**

**就发现95%的这个议题**

**are quite consistent,**

**其实都是比较一致的**

**for instance,**

**比如说**

**human-centered approach**

**human-centered approach**

**以人类为中心也好**

**以人类为本也好**

**Then, for example,**

**然后再比如说**

**ensuring that AI systems are transparent and explainable,**

**要保证人工智能系统是透明 可解释的**

**ensuring the safety of AI systems,**

**要维护人工智能系统的安全**

**ensuring fairness,**

**要具备公平性**

**and protecting privacy.**

**保护隐私**

**These topics I think**

**这些议题 我想**

**are globally recognized as consensual.**

**全球可以说它是有共识的**

**So, in 2021,**

**所以在2021年的时候**

**the UNESCO Recommendation on the Ethics of AI was adopted.**

**就是联合国教科文组织就推出了**

**It was unanimously adopted by all its 193 member states.**

**全球193个会员国达成一致的《人工智能伦理问题建议书》**

**At that time, our country's**

**我们国家**

**permanent representative to UNESCO**

**当时我们在教科文的常驻代表**

**signed it on behalf of China.**

**也代表中国在上面签了字**

**We can say that the UNESCO's**

**可以说教科文组织的**

**Recommendation on the Ethics of AI**

**这一份《人工智能伦理问题建议书》**

**has already established a global consensus on AI ethics.**

**在人工智能伦理方面 已经形成了全球的共识**

**On the other hand,**

**那么另外一方面就是**

**what is the difference between our vision and others'?**

**我们的愿景跟别人的愿景有什么不同**

**From 2017 to 2018, we advocated**

**2017年 2018年开始倡导**

**for ethical AI.**

**合乎伦理的人工智能**

**From 2019 to 2022, we talked about AI governance.**

**2019年到2022年讲人工智能的治理**

**And from 2023 to 2024,**

**然后到2023年 2024年**

**we are focusing on AI safety issues.**

**聚焦在人工智能的安全的这个问题**

**We can say that**

**我们可以说**

**the global governance of AI**

**在人工智能的全球治理的方面**

**is becoming increasingly focused**

**这个议题 也是越来越聚焦**

**and concrete.**

**越来越落地**

**An important event in 2023 was**

**那么在2023年 发生的比较重要的事情**

**28 countries plus the European Union**

**就是28国加上欧盟**

**signing the Bletchley Declaration on AI Safety**

**签署了人工智能安全的**

**in Bletchley, the UK.**

**在英国的布莱切利《布莱切利宣言》**

**This marked the first time**

**那么这是全世界第一次**

**the world had reached a resolution**

**针对人工智能安全的这个议题**

**with wide consensus**

**形成了一个具备一定广泛性**

**on the issue of AI safety.**

**这样的一个决议**

**However,**

**当然了**

**it was not a resolution at the United Nations level.**

**它还不是联合国层面的决议**

**So there is a call for establishing**

**所以后来有很多的号召说**

**a specialized international organization for AI,**

**人工智能应当成立专门的国际组织**

**similar to the International Atomic Energy Agency.**

**就像国际原子能委员会一样**

**成立一个专门的国际组织**

**The UN Secretary-General**

**那么联合国秘书长**

**responded to this call for**

**也响应这样的号召**

**establishing a high-level advisory body on AI.**

**专门成立了人工智能高层的顾问机构**

**So it was discussed**

**那么去探讨就是**

**whether the UN should**

**联合国是不是应该需要**

**establish a special agency for AI.**

**成立一个专门的这个机构**

**Following this, China took several actions concerning**

**比如说后来中国针对**

**the UN General Assembly resolution**

**人工智能能力建设的**

**on AI capacity building.**

**这项联大的决议**

**后来我们就做了好几件事**

**First,**

**第一 我们在上海**

**with the support of the UN,**

**在联合国的支持下**

**we hosted the first seminar**

**召开了第一届**

**on AI capacity building in Shanghai.**

**人工智能能力建设的研讨班**

**Representatives from 40 countries**

**来自40个国家的代表**

**participated in discussions on AI capacity building.**

**来参与人工智能能力建设的研讨**

**This was China's**

**那么人工智能能力建设研讨班**

**first attempt at cooperation**

**是我们国家做的第一个尝试**

**between the Ministry of Foreign Affairs and the UN.**

**是外交部和联合国的合作**

**A recent development is**

**那么第二就是近期**

**the establishment of**

**在联合国成立了**

**a group of friends for international cooperation**

**on AI capacity building within the UN,**

**人工智能能力建设国际合作之友小组**

**which is led primarily by China,**

**这个也是由我们国家为主**

**with 80 participating countries.**

**80国参与的**

**The formation of this group signifies**

**成立之友小组什么意思**

**organized efforts towards**

**就是你做全球人工智能能力建设**

**global AI capacity building.**

**这件事情有组织了**

**Under the umbrella of the UN**

**在联合国的组织下**

**and coordination by this group,**

**在之友小组的这种协调下**

**countries can**

**各个国家**

**genuinely advance**

**就真正能推进**

**their AI capabilities.**

**这些人工智能能力的建设**

**I think these are just the beginning.**

**但我想这些仍然是一些开始**

**Professor Gao, what's your view?**

**高老师**

**I think the issue of AI governance**

**人工智能治理问题**

**embodies the concept of**

**我觉得它更体现**

**a community with a shared future for mankind.**

**人类命运共同体这个理念**

**The risks posed by AI**

**因为人工智能带来的风险**

**are systemic and global;**

**它是一个系统性的 全球性的**

**they are by no means confined to any single country.**

**它绝不是一个国家的**

**To achieve global governance of AI,**

**那么人工智能达到全球性治理**

**the first step is to**

**第一步**

**transform these principles into rules.**

**要把原则变成规则化**

**This means that each country should,**

**规则化就是说要成为各个国家**

**through its sovereign legislation,**

**各自通过主权国家的立法**

**turn them into legal regulations.**

**把它变为法律规则**

**When turning these principles into legal rules,**

**那么在变成法律规则的时候**

**I personally believe that**

**我个人觉得**

**legislation across the world should**

**世界范围内各国家的立法**

**move towards standardization.**

**都应当朝着规范化的方向发展**

**This means that the rules different countries formulate**

**你制定的那个规则和我制定规则**

**should be consistent in standards.**

**咱们的规范应当是一致的**

**However,**

**但是现在由于各个国家**

**due to different legislative traditions**

**它的立法的传统不一样**

**and focuses among countries,**

**着眼点不一样**

**achieving standardization alone is not enough.**

**光规范这一步我觉得还不够**

**Going forward,**

**就是再往下**

**it may be necessary to convert the legal norms**

**可能要把各国家自己的**

**formulated by each country based on international principles**

**根据国际原则制定的法律规范**

**into a set of**

**你要把它转变成**

**systems or codes**

**计算机可识别的一套系统**

**recognizable by computers.**

**一套代码**

**Only in this way**

**只有在这种情况下**

**can AI develop in accordance with**

**那么人工智能**

**the consensus reached by humanity,**

**才能够沿着我们人类的指引达成这些共识**

**moving towards a benevolent and friendly direction.**

**向着向善友好的方向来发展**

**I strongly disagree with the notion that**

**我特别不赞成说**

**developing AI**

**我们搞了人工智能**

**should be about**

**就要维持谁的**

**maintaining someone's status in the world.**

**在世界上的什么地位**

**I think this is incorrect**

**我觉得这是不对的**

**because it is fundamentally an issue concerning all of humanity.**

**因为它本身它是一个全人类的问题**

**It is not a problem specific to one country.**

**它不是一个国家的问题**

**If the development of AI**

**如果只是说**

**is solely aimed at**

**我发展人工智能**

**preserving hegemony in a certain field,**

**是为了保持在某个领域的霸权地位**

**that would be entirely wrong.**

**那就完全错误了**

**Such an approach could lead to**

**那就成了真是一个不可想象的**

**unimaginable consequences.**

**一个后果了**

**At the end of the program,**

**那么在节目的尾声**

**I would like to ask both experts to**

**我们请二位专家**

**summarize your viewpoints**

**各用一段话总结一下**

**discussed in the show**

**今天我们在节目当中输出的观点**

**and also your feelings.**

**和大家参加节目的感受**

**Professor Zeng, please.**

**我想请曾老师先来**

**I remember**

**我还是想起我当时做**

**when I was working on UNESCO's Recommendations**

**联合国教科文组织的**

**on the Ethics of AI,**

**《人工智能伦理问题建议书》当中**

**there was a section where**

**有一个环节**

**I wanted to incorporate the idea of harmonious coexistence**

**我希望把和谐共生**

**into the ethical principles of AI.**

**然后写到人工智能的伦理原则当中**

**In the final document, it was articulated as peaceful coexistence.**

**最后写进去的方式叫做和平共生了**

**Coexistence can also be translated into**

**共生在英文里翻译成**

**"optimizing symbiosis."**

**optimizing symbiosis**

**It means that through collective efforts on a global scale,**

**它的意思是全球通过共同的努力**

**the world will become better and better.**

**世界会变得越来越好**

**This is not just about peace;**

**那么它不仅仅是和平**

**it must involve a virtuous developmental process.**

**它要有一个向善的发展的过程**

**Our concept of "promoting global solidarity"**

**那么现在我们倡导同球共济**

**is essentially about how AI,**

**是我们未来人工智能实际上**

**as a widely empowering technology,**

**它是作为一种广泛的赋能技术**

**can genuinely promote the building of**

**就真正地推进**

**a community with a shared future for mankind.**

**全球人类命运共同体的实现**

**I believe this is the underlying philosophy here.**

**我想背后的理念是这样的**

**Now, Professor Gao please.**

**高老师**

**As a legal professional, my view is that**

**我作为一个法律工作者**

**when it comes to legal norms**

**那么法律规范**

**in relation to a technology-driven field,**

**在面对人工智能**

**such as AI,**

**这样一个高度依赖技术的一类对象**

**we need to have standardization of legal norms**

**可能还要实现的法律规范要标准化**

**and establish technical standards.**

**要达成许多技术标准**

**Only after we achieve a common set of rules**

**我们实现规则**

**and standards**

**标准共同化以后**

**can we code these rules into practice,**

**我们才能够实现规则标准的代码化**

**ultimately working towards a form of AI**

**最后可能我们实现人工智能的**

**that is good, friendly, and automatic,**

**向善 友好 自动执行**

**thereby realizing its intelligence.**

**才能实现它的智能化**

**That's my vision of future AI governance.**

**所以我想了“五化”的一个愿景**

**The whole world should work together in this direction.**

**未来应当全世界共同朝这个方向去努力**

**As the host,**

**我作为主持人**

**what I've learned from both the experts today**

**今天听二位专家分享的观点 从中学到的是**

**is that the development of AI**

**人工智能它的发展**

**does not have minor premises,**

**没有小前提**

**only major premises.**

**只有大前提**

**The essence of China's proposal for**

**我们国家提出来的**

**responsible AI development**

**负责任地发展人工智能**

**is not to serve an individual**

**本质上的目的**

**is not to serve an individual**

**并不是为了服务一个“小我”**

**but to serve the collective,**

**而是为了服务一个“大我”**

**which refers to our planet.**

**这个“大我”是指我们这颗蓝色的星球**

**It is about**

**这颗智慧星球上的**

**all countries,**

**所有的国家**

**all ethnic groups,**

**所有的人种**

**and all the people on this smart planet.**

**所有的人**

**I hope that AI can enable everyone**

**希望人工智能能够让所有的人**

**to live a good life.**

**都过上好日子**

**This is also what we refer to as "promoting global solidarity".**

**这也是我们说的同球共济**

**That's all for this episode of TechWorld.**

**好的这里是节目《科技π》**

**I am your host, Wu Baojun.**

**我是主持人吴宝俊**

**If you have enjoyed our program, please follow us.**

**喜欢我们节目的朋友可以长期关注**

**See you next time.**

**咱们下期再见**