联合国 S/PV.9381

安全理事会

第七十八年

临时逐字记录

第九三八一次会议

2023年7月18日星期二上午10时举行

纽约

主席:	克莱弗利先生/吴百纳女爵士	(大不列颠及北爱尔兰联合王国)
成员:	阿尔巴尼亚	霍查先生
	巴西	弗兰萨·丹尼斯先生
	中国	张军先生
	厄瓜多尔	佩雷斯·卢塞先生
	法国	德里维埃先生
	加蓬	恩格耶马·恩东夫人
	加纳	阿杰曼先生
	日本	武井先生
	马耳他	弗雷泽夫人
	莫桑比克	贡萨尔维斯先生
	俄罗斯联邦	波利扬斯基先生
	瑞士	贝里斯维尔夫人
	阿拉伯联合酋长国	谢拉夫先生
	美利坚合众国	德劳伦蒂斯先生

议程项目

维护国际和平与安全

人工智能: 国际和平与安全面临的机遇与风险

23-21048 (C)

上午10时05分开会。

通过议程

议程通过。

维护国际和平与安全

人工智能: 国际和平与安全面临的机遇与风险

主席(以英语发言): 我热烈欢迎秘书长和各位高级代表。他们今天与会凸显了所讨论主题的重要性。

根据安理会暂行议事规则第39条, 我邀请下列通报人参加本次会议: Anthropic公司联合创始人杰克·克拉克先生和中国科学院自动化研究所的曾毅先生。

主席(以英语发言): 安全理事会现在开始审议其议程上的项目。

我谨提请安理会成员注意文件S/2023/528, 其中载有2023年7月14日大不列颠及北爱尔兰联合王国常驻联合国代表给秘书长的信, 转递一份关于所审议项目的概念说明。

我现在请秘书长安东尼奥·古特雷斯阁下发言。

秘书长(以英语发言): 我感谢联合王国召开安全理事会有史以来第一次关于人工智能的辩论会。

我关注人工智能的发展已有一段时间。事实上, 我六年前就告诉大会, 人工智能将对可持续发展、工作领域和社会结构产生巨大影响。不过, 和在座各位一样, 我对最新形式的人工智能——生成式人工智能——感到震惊和印象深刻, 因为它在能力上取得了根本性进步。无论其形式为何, 这一新技术的速度和影响都是完全史无前例的。有人把它与印刷机的问世相比。然而, 印刷书籍在欧洲的普及用了50多年时间, 而 ChatGPT 在短短两个月内就拥有了 一亿用户。据金融业估计, 到2030年, 人工智能将为全球经济贡献10至15万亿美元。世界上几乎每个国家的政府、每个大公司和组织都在制定人工智能战略。但是, 即使是它自己的设计师也不知道他们令人惊叹的技术突破可能会带来什么结果。

很明显, 人工智能将对我们生活的每个领域产生影响, 其中包括联合国的三大支柱。从监测气候危机到医学研究的突破, 它有可能推动全球发展。它为享有人权, 特别是卫生和教育方面的人权提供了新的可能。但是, 有证据表明, 人工智能可能放大偏见, 强化歧视, 并使专制政府能进一步加强监控, 人权事务高级专员已对此表示震惊。

在人工智能在和平与安全领域已经引起政治、法律、道德和人道主义方面关切的情况下, 今天的辩论为审议人工智能对和平与安全的影响提供了机会。我敦促安理会带着紧迫感和全球视角, 以学习者的心态对待这一技术, 因为我们所看到的只是一个开始。技术创新的脚步再也不会像今天这样缓慢。

人工智能正在被用于和平与安全方面的工作, 这样做的包括联合国。人工智能正越来越多地被用于查明暴力模式、监测停火等, 帮助加强我们的维和行动、调解和人道主义努力。但是, 人工智能工具也可能被那些怀有恶意的人使用。人工智能模型可以帮助人们大规模地伤害自己和彼此。

我们应清楚一点: 利用人工智能系统从事恐怖和犯罪活动, 或达成国家目的, 此种恶意之举可能会导致可怕的死亡和破坏、广泛的创伤以及难以想象的深度心理伤害。人工智能支持的网络攻击已在把关键基础设施及我们自己的维和与人道主义行动作为目标, 给人类造成巨大的痛苦。利用人工智能的技术和财政障碍很低, 包括对犯罪分子和恐怖分子而言。

人工智能的军事和非军事应用都可能给全球和平与安全造成非常严重的后果。生成式人工智能的出现可能会成为虚假信息和仇恨言论的决定性时刻, 它破坏真相、事实和安全, 为操纵人类行为增添新的维度, 大大加剧两极分化和不稳定。深伪技术只是一种新的人工智能驱动的工具, 如果不加以控制, 可能会对和平与稳定造成严重影响。一些人工智能支持的系统影响尚难预料, 可能会意外地造成安全风险。

看看社交媒体就知道了。原本旨在加强人类联系

的工具和平台现在被用来破坏选举，传播阴谋论，煽动仇恨和暴力。人工智能系统运转失灵是另一个值得关注的重大领域。人工智能与核武器、生物技术、神经技术和机器人技术之间的相互作用令人深感震惊。

生成式人工智能在大规模行善或作恶方面潜力巨大。它的创作者自己也警告说，未来将面临更大、有可能是灾难性的生存风险。如果不采取行动应对这些风险，我们就没有尽到对今世后代的责任。

（以法语发言）

国际社会在应对有可能破坏我们社会和经济的新技术方面有着悠久的历史。我们曾在联合国汇聚一堂，制定新的国际规则，签署新的条约，建立新的全球机构。许多国家呼吁围绕人工智能治理采取不同的措施和举措，但这需要一种普遍的做法。治理问题将是复杂的，原因有几个。

第一，强大的人工智能模型已经可以为普通大众广泛应用。第二，与核材料和生化制剂不同，人工智能工具可在世界各地移动，几乎不会留下任何痕迹。第三，私营部门在人工智能领域的主导作用在其他战略技术领域几乎没有可比性。

不过，我们已经找到切入点。一个是通过《特定常规武器公约》达成的2018-2019年致命自主武器系统指导原则。许多专家建议禁止无人控制的致命自主武器，我对此表示赞同。另一个是通过教科文组织商定的2021年人工智能伦理建议。反恐怖主义办公室与区域间犯罪和司法研究所合作，就会员国如何应对人工智能可能被用于恐怖主义目的的问题提出了建议。由国际电信联盟主办的“人工智能惠及人类”峰会将专家、私营部门、联合国机构和各国政府聚集在一起，共同努力确保人工智能为公益服务。

（以英语发言）

最好的办法是应对现有挑战，同时建立监测和化解未来风险的能力。它应具有灵活性和适应性，并考虑到技术、社会和法律问题。它应把私营部门、民间社会、独立科学家和人工智能创新的所有推动者整合在一起。制定全球标准和办法的必要性使联合国成为实现这一目标的理想场所。《宪章》强调保护子孙后代，这给了我们一项明确的任务，即把所有利益攸关方聚集在一起，共同减轻长期全球风险。人工智能就是这样一种风险。

因此，我欢迎一些会员国呼吁在国际原子能机构、国际民用航空组织和政府间气候变化专门委员会等模式的启发下，成立一个新的联合国实体，为监管这项非凡技术的集体努力提供支持。该机构的首要目标将是支持各国最大限度利用人工智能带来的益处，从事有益的活动，减轻现有和潜在的风险，以及建立和管理国际商定的监测和治理机制。

实事求是地说：政府和其他行政及安全机构在人工智能方面存在着巨大的技能差距，必须在国家和全球层面加以解决。一个新的联合国实体将收集专门知识，供国际社会使用。它还可以支持在人工智能工具的研究和开发方面的协作，为可持续发展提速。作为第一步，我正在召集一个多利益攸关方人工智能高级别咨询委员会，该委员会将在今年年底前就全球人工智能治理的备选方案提出报告。我即将就《新和平纲领》提供的政策简报还将就人工智能治理向会员国提出建议。

第一，它将建议会员国根据国际人道法和人权法规定的义务，制定负责任地设计、开发和使用人工智能的国家战略。

第二，它将呼吁会员国参与多边进程，围绕人工智能的军事应用制定规范、规则和原则，同时确保其他相关利益攸关方的参与。

第三，它将呼吁会员国商定一个全球框架，以规范和加强对把包括人工智能在内的数据驱动技术用于反恐的做法进行监督的机制。

关于《新和平纲领》的政策简报还将呼吁在2026年之前结束谈判，制定一项具有法律约束力的文书，禁止可在无人控制或照看的情况下运作且不能按照国际人道法使用的致命自主武器系统。

我希望, 会员国将对这些备选办法进行辩论, 并就建立急需的人工智能治理机制的最佳行动方针作出决定。

除了在《新和平纲领》中提出建议外, 我敦促各国就这样一个总原则达成一致意见：即人的操纵和控制对核武器而言不可或缺, 而且永远都不应取消。将于明年举行的未来峰会将是就许多此类相互关联的问题作出决定的理想机会。

我敦促安理会在人工智能领域发挥领导作用, 并展示我们可以采取的共同措施, 以确保人工智能系统的透明度、问责制和监督。我们必须共同努力, 让人工智能能够弥合社会、数字和经济隔阂, 而不是让我们进一步分裂。我敦促安理会成员齐心协力, 为和平与安全建立信任。我们需要一场开发人工智能造福人类的竞赛, 这样的人工智能应该是可靠而安全的, 可以终结贫困、消除饥饿、治愈癌症、驱动气候行动、推动我们实现可持续发展目标。这才是我们需要的竞赛, 这种竞赛是可能的, 也是可以实现的。

主席（*以英语发言*）：我感谢秘书长的通报。

我现在请克拉克先生发言。

克拉克先生（*以英语发言*）：我今天在这里简要介绍为什么人工智能已经成为世界各国关注的主题, 未来几年该技术的发展如何, 谈谈政策制定者可以如何选择应对这一历史机遇的一些想法。我发言的要点应该是, 我们不能把人工智能的开发完全交给私人部门行为体来做。世界各国政府必须团结起来, 发展国家能力, 让强大的人工智能系统的开发成为社会各界的共同努力, 而不是仅仅由少数在市场上相互竞争的公司来决定。

我为什么要说这番话? 它有助于理解最近这段历史。十年前, 英格兰一家名为DeepMind的公司发表了一项研究, 展示了如何教人工智能系统游玩一些计算机老游戏, 如《太空侵略者》和《乓》。快进到2023年, 那项研究中使用的同样一些技术现在正被用于创建人工智能系统, 它们可以在空战模拟中击败军事飞行员, 稳定聚变反应堆中的等离子体, 甚至设计下一代半导体的组件。类似的趋势也在计算机视觉领域上演。十年前, 科学家能够创建基本的图像分类器, 并生成非常粗糙的像素化图像。如今, 图像分类在世界各地被用于检查生产线上的产品, 分析卫星图像和提高国家安全。当今引起关注的人工智能模型, 如OpenAI的ChatGPT、谷歌的Bard和我自己的公司Anthropic的Claude, 它们本身也是出于企业利益开发的。因此, 10年来发生了很多事, 我们可以预计未来几年出现新的、更强大的系统。我们可以预计这些趋势将继续下去。在世界范围内, 私人部门行为体拥有先进的计算机以及大量的数据和资本资源来构建这样的系统, 因此私人部门行为体看上去很可能继续决定着这种系统的发展。虽然这将给全世界人类带来巨大好处, 但也对和平、安全和全球稳定构成潜在威胁。

这些威胁源于人工智能系统的两个基本特性。第一是它们被滥用的可能性,第二是它们的不可预测性, 以及它们系由一批如此有限的行为体开发这一事实本身所包含的脆弱性。关于滥用, 人工智能系统具有越来越广泛的能力, 一些有益的能力与其他能力并存, 可能构成严重滥用的威胁。例如, 可以帮助我们理解生物学的人工智能系统也可以用来制造生物武器。关于不可预测性, 人们对人工智能的一种基本感觉是, 我们不理解它的系统。这就好像我们不理解燃烧的科学原理, 就制造了发动机。这意味着, 一旦人工智能系统被开发和部署, 人们就会为它们找到开发者没有预料到的新用途。其中许多用途将是积极的, 但正如我之前提到, 有些可能是滥用行为。行为混乱或不可预测的问题更具挑战性。人工智能系统一旦部署, 就可能会表现出在开发过程中没有发现的微妙问题。因此, 我们应该非常仔细地思考如何确保对这种系统的开发者进行问责, 从而要求他们建立和部署不会危及全球安全的安全可靠的系统。

为了强调这个问题, 或许可以用一个类比来帮助理解。我想要求每一个聆听这一通报的人不要把人工智能看作一种特定的技术, 而是一种人力劳动, 它可以以计算机的速度来买卖, 并且随着时间推移而变得

越来越廉价且更有能力。正如我描述的那样，这是一种由一类有限的行为体——公司——所发展的劳动形式。我们应该看清它能提供的巨大政治影响力。如果你能创造出人力劳动的替代品或增强物，并将其出售给全世界，那么你将逐渐变得更有影响力。如果我们用这种方式来看待人工智能，似乎更容易思考人工智能政策的许多挑战。任何人只要有足够的金钱和数据，就可以在某个特定领域轻松创造出一个人造专家，对此,世界各国应该作何反应？谁应该获得这种权力？政府应该如何监管这种权力？谁能够成为可以创造和贩售这种所谓专家的行为体？我们可以被允许创造哪些类别的专家？这些都是重大问题。

根据我的经验，我认为我们可以做的一件有用的事情是努力开发各种方法来测试这些系统中的能力、滥用情况和安全隐患。如果我们创造和分配的是即将进入全球经济的新型工作者，就理应能够描述其特征、评估其能力并了解其缺陷。毕竟，从急救服务到军事领域，人类在许多关键岗位上都要经过严格评估和在职测试。人工智能为什么不用这样做？

因此，令人鼓舞的是，许多国家在其各种人工智能政策提案中强调了安全测试和评估的重要性，从欧盟的人工智能框架到中国最近宣布的生成式人工智能规则，从美国国家标准与技术研究院的人工智能系统风险管理框架到即将举行的人工智能与人工智能安全伦敦峰会。所有这些不同的人工智能政策提案和活动都以某种形式依赖于对人工智能系统的测试和评估，以便让世界各国政府可以在该领域进行投资。目前，对于如何测试前沿系统的歧视、滥用和安全等问题，还没有标准甚至最佳做法。而因为没有最佳做法，政府很难制定政策，对开发系统的行为体进行更大程度的问责。相应地，这意味着私营部门行为体在与政府交涉时享有信息优势。

总之，任何明智的监管方法首先都要有能力评估人工智能系统的特定能力或缺陷。任何失败的方法首先都是只有宏大的政策构想，却没有实际的衡量和评估来支持。要通过开发强大、可靠的评估系统，政府才能够让公司承担责任，公司才能够赢得它们希望部署人工智能系统的这个世界的信任。如果我们不投资于此，就会存在“监管俘虏”的风险，危及全球安全，并将未来交到一小批私人部门行为体的手中。然而，如果我们能够迎难而上，我相信我们可以获得人工智能作为一个全球社群所带来的好处，并确保人工智能开发者和世界公民之间的权力平衡。

主席（以英语发言）：我感谢克拉克先生所作通报。

我现在请曾毅先生发言。

曾毅先生（以英语发言）：我叫曾毅。我要借此机会与安理会分享我个人对人工智能的看法，我认为它是促进国际和平与安全的一股向善力量。我希望它将有助于促进对人工智能全球治理必要性的讨论和理解。

毫无疑问，人工智能是推进全球可持续发展的强大赋能技术。在调查人工智能在可持续发展目标方面的应用时，我们发现大多数努力都集中在人工智能促进优质教育和医疗保健方面，而在许多其他重要问题上，如人工智能促进生物多样性、气候行动或和平等方面，却很少有人作出努力。然而，我认为这些都是对人类未来至关重要的问题，各国政府当然应该共同努力解决这些问题。

军事人工智能与用于和平与安全的人工智能虽然密切相关，但在一些重要方面有根本区别。我们应当推动使用人工智能促进国际和平，将之作为可持续发展目标的重要支柱，以减少而不是增加安保和安全风险。当我们从和平与安全的角度思考人工智能的好处时，与其设法为军事和政治目的制造虚假信息，不如致力于使用人工智能来识别不同国家和政治机构之间的虚假信息和误解，并使用它来保卫网络而不是攻击网络。

人工智能应当用来连接人民和文化，而不是切断这种联系。因此，我们创建了一个人工智能支持的文化互动引擎，在各个教科文组织世界遗产地之间寻找共同点和多样性。这些世界遗产地表明，我们在文化方面并不像我们想象的那样充满分歧，它们之间的共

同点就像长出的根和茎一样，帮助我们欣赏、理解各种文化的多样性，甚至从中学习。

目前人工智能的例子，包括最近的生成式人工智能，都是看似智能但没有真正理解能力的信息处理工具，因而并不是真正的智能。当然，也正是因此，不能信任它们作为负责任的代理者来帮助人类作决定。例如，尽管世界尚未就致命自主武器系统达成足够广泛的共识，但至少有一点是明确的，即不应直接使用人工智能来为人类作出生死攸关的决定。为了确保人类与人工智能的适当互动，必须应用有效和负责任的人类控制。人工智能也不应用于使外交任务自动化，特别是国家之间的外交谈判，因为它可能会利用和放大欺骗和不信任等人类局限性和弱点，给人类带来更大的、甚至灾难性的风险。人工智能模型中没有“我”、甚至“自我”的概念，而由生成式人工智能驱动的对话系统却总是在论证中使用“我认为”和“我建议”这样的表达方式，这是奇怪的、误导性的，甚至是不负责任的。因此，我要再次强调，人工智能永远不应假扮人类，取代人类或误导人类产生错误的认知。我们应当使用生成式人工智能来帮助我们，但绝不要相信它能取代人类的决策。

我们必须确保所有人工智能武器系统由人类控制，而且人类的控制必须是充分、有效和负责任的。例如，必须避免人类与人工智能互动过程中的认知过载。我们必须防止人工智能武器系统的扩散，因为相关技术很可能被滥用或恶意使用。无论是短期还是长期，人工智能都有可能导致人类灭绝，原因很简单，我们尚未找到一种方法来保护自己，避免人工智能利用人类的弱点——如果人工智能这样做了，它甚至不知道我们所说的人类、死亡和生命是什么意思。从长远来看，我们还没有给超级智能提出它应该保护人类的任何切实理由，而解决这个问题的办法可能需要几十年才能找到。我们的初步研究表明，我们可能必须改变彼此之间以及与其他物种、生态和环境之间的互动方式。这可能需要全人类共同作出决策，我们都必须在这方面深入合作，深度思考。

考虑到这些短期和长期挑战，我很肯定我们今天还无法解决人工智能促进和平与安全的问题，但本次具有挑战性的讨论，对会员国来说仍可能是一个良好的起点。在这方面，我谨建议安全理事会考虑是否有可能设立一个人工智能促进和平与安全工作组，就短期和长期挑战开展工作，因为在专家一级开展合作将更加灵活和科学，更容易从科学和技术角度达成共识，并为安理会成员的决策提供协助和支持。安理会成员应当为其他国家树立良好榜样，并在这一重要问题上发挥重要作用。

人工智能应当帮助人类解决问题，而不是制造问题。曾经有一个男孩问我，人工智能辅助的核弹除了是科幻小说中的一个特色外，可否用来摧毁攻击地球的小行星或改变其轨道，以防止与地球相撞，从而拯救我们的生命。虽然这个想法在科学上可能不太靠谱，目前来看可能非常危险，但它至少是使用人工智能来解决人类的问题，总比让人工智能帮助我们在地球上用核武器互相攻击要好得多，那样做会给人类社会带来问题，并可能给我们、下一代甚至整个人类文明带来灾难性风险。在我看来，人类应当始终保持对使用核武器的最终决策权并对此负责。我们已经申明，核战争打不赢也打不得。许多国家——包括但不限于安理会五个常任理事国——已经宣布各自关于将人工智能用于安全以及总体治理的战略和观点，我们可以看到有一些共同点，可以作为国际共识的重要投入，但这仍然不够。联合国必须在建立人工智能促进发展和治理的框架方面发挥核心作用，以确保全球和平与安全。在人类命运共同体的背景下，我们必须共同制定这一议程和框架，不让任何人掉队。

主席（以英语发言）：我感谢曾毅先生所作的通报。

我现在以联合王国代表的身份发言。

这是安全理事会首次讨论人工智能，是一次历史性会议。自艾伦·图灵和克里斯托弗·斯特雷奇等先驱对人工智能进行早期开发以来，这项技术的发展速度越来越快。然而，人工智能引发的最大变革还在后头。虽然我们无法完全理解这些变革的规模，但人

类的收益肯定是巨大的。人工智能将从根本上改变人类生活的方方面面。医学上的突破性发现可能即将到来。我们经济的生产力提升可能是巨大的。人工智能可以帮助我们适应气候变化、打击腐败、革新教育、实现可持续发展目标并减少暴力冲突。

但我们今天在这里开会是因为人工智能会影响安理会的工作。它可能加强，也可能破坏全球战略稳定。它对我们关于防御和威慑的基本假设构成挑战。它对追究战场上致命决定的责任提出道德问题。毫无疑问，人工智能正在改变虚假信息的速度、规模和传播，给民主和稳定带来巨大的有害后果。

人工智能可以帮助国家和非国家行为体不计后果地寻求获得大规模毁灭性武器，但也可以帮助我们阻止扩散。这就是为什么我们迫切需要建立对变革性技术的全球治理——因为人工智能没有国界。

联合王国的愿景是建立在四项不可减少的原则之上的。它应该是开放的-人工智能应该支持自由和民主。它应该负责任-人工智能应该与法治和人权保持一致。它应该是安全的-人工智能应该在设计上是安全、可预测的，保护产权、隐私和国家安全。它应该具有弹性-人工智能应该得到公众的信任，关键系统必须得到保护。

联合王国的做法建立在现有多边倡议的基础上，如日内瓦的人工智能造福人类全球峰会，或教科文组织、经济合作与发展组织和20国集团的工作。人工智能全球伙伴关系、七国集团的广岛人工智能进程、欧洲委员会和国际电信联盟等机构都是重要的合作伙伴。开拓性的人工智能公司也需要与我们合作，以便我们能够获得收益，并将人类面临的风险降至最低。没有一个国家能够免受人工智能的影响，因此我们必须让各个领域的国际行为体参与并加入，结成最广泛的联盟。

联合王国是世界上许多开拓性人工智能开发人员和最重要的人工智能安全研究人员的家园。这就是为什么今年秋天，联合王国计划将世界各国领导人聚集在一起，举行首次关于人工智能安全的重大全球峰会。我们的共同目标是审视人工智能的风险，并决定如何通过协调行动来降低这些风险。

各种机遇摆在我们面前，其规模之大，我们几乎无法想象。我们必须抓住这些机遇，从全球在基本原则上团结一致这一立场出发，果断、乐观地应对人工智能的挑战，包括对国际和平与安全的挑战。

“人生世事如潮汐，若能把握机会乘风破浪，必定能马到成功。”本着这种精神，让我们共同努力，在我们跨过一个陌生世界的门槛时确保和平与安全。

我现在恢复履行安理会主席的职责。

我现在请日本外务副大臣发言。

竹井先生（日本）（*以英语发言*）：我赞扬联合王国倡议在安全理事会讨论人工智能问题。这将是未来全球讨论的良好开端。我也感谢秘书长和其他通报者。

人工智能正在改变世界。人工智能改变了人类的生活。它的速度、潜力和风险超出了我们的想象，超越了国界。我们正在这个历史关头受到考验：我们能否自律来控制它？

我的政治信念是“与其担心，不如应对”。我认为，应对挑战的关键有两个：人工智能必须以人为本，必须值得信赖。人类可以也应该控制人工智能来增强人类潜力，而不是相反。让我提出两点。

第一，关于以人为本的人工智能，人工智能的发展应该与我们的民主价值观和基本人权保持一致。人工智能不应成为统治者的工具，而应置于法治之下。人工智能在军事上的应用就是一个很好的例子。此种应用应该负责任、透明并以国际法为基础。因此，日本将继续在《特定常规武器公约》的框架内为关于致命自主武器系统的国际规则制定进程作出贡献。

其次，关于值得信赖的人工智能，如果让广泛的利益攸关方参与规则制定进程，人工智能可以更加值得信赖。我认为，在这方面，联合国的召集力能够发挥作用，汇聚世界各地的智慧。上个月，日本与反恐怖

主义办公室和联合国区域间犯罪和司法研究所共同举办了一场会外活动， 在联合国主导了关于恐怖分子滥用人工智能问题的讨论。日本还感到自豪的是，今年启动了七国集团广岛人工智能进程，并为全球关于生成式人工智能的讨论作出了贡献。

安全理事会和联合国可以通过使用人工智能来更新其现有工具包。首先，我们应该考虑如何积极使用人工智能提高安理会决策和工作方法的效率和透明度。我们欢迎秘书处努力利用人工智能开展调解和建设和平活动。此外，我们可以采用基于人工智能的冲突预警系统、制裁执行监测以及打击有关和平行动的虚假信息的反措施，使联合国的工作更有效率和效益。

最后，请允许我表示，我们愿意积极参加联合国内外关于人工智能的讨论。

主席（*以英语发言*）：我现在请莫桑比克共和国外交与合作部副部长发言。

贡萨尔维斯先生（莫桑比克）（*以英语发言*）：莫桑比克热烈祝贺联合王国出色地担任主席，并召开今天关于人工智能的及时而重要的辩论会，重点是人工智能给国际和平与安全带来的机遇与风险。

我们谨感谢联合国秘书长安东尼奥·古特雷斯先生阁下富有洞察力的发言。我们感谢通报者所作的深思熟虑且中肯的发言。

首先，请允许我坦诚地披露：这一发言完全是人工写成的，而不是由诸如著名的ChatGPT这样的生成式人工智能模型做出的。这一披露很重要。它揭示了围绕人工智能快速进步的一些焦虑，特别是我们正在接近这样一个节点：数字机器可以执行那些在人类存在的大部分时间里完全属于人类智能领域的任务。

最近，人工智能系统的能力和可见度都在加速，同时人们对其能力和局限性的认识也在不断增强，这引发了人们的担忧,即这项技术的发展速度如此之快，可能不再安全可控。这种谨慎是必要的。

虽然人工智能的最新进展为通过创新的民主化来增强演讲、医学和战争等不同领域提供巨大的机会，但某些模型也已展示超出其创造者理解和控制的能力。这带来各种各样的风险，包括产生灾难性后果的可能性。事实上，我们应该注意“魔术新手”的警示故事。

随着人工智能引擎越来越具有令人信服的模仿力，在某些情况下甚至超越与人类相关的各种行为，它们成为传播错误信息、诈骗个人、从事学术作弊、欺骗性引发冲突、招募恐怖分子、散布分裂以及实施许多其他邪恶活动的理想工具。

人工智能模型已发展成为自我编程机器，能够通过循环往复的自我改进来使其学习过程自动化。因此，必须建立强有力的治理结构，以期减少事故和滥用的风险，同时仍然促进创新并利用取得积极成果的潜力。

正如本次通报会的概念说明所准确指出的那样，人工智能技术具有深刻改变我们各社会，产生大量积极效果的潜力。人工智能有助于消除疾病，应对气候变化并准确预测自然灾害，从而使之成为全球南方的宝贵盟友。

同样，通过利用坚持质量控制、获取以及会员国参与等严格标准的组织，如社会冲突分析数据库、区域机构和整个联合国系统所建立的广泛数据库，我们有可能加强预警能力、规划调解努力，并加强维持和平方面的战略沟通等等。人工智能可以成为一个宝贵的工具，可以利用这些海量数据来为这些努力服务。

面对使用人工智能所带来的机遇和威胁，作为联合国会员国，莫桑比克共和国认识到，应当采取包含以下方面的平衡办法。

首先，一旦出现可信的证据表明人工智能构成生存风险，那就必须谈判一项政府间条约，来管理和监督人工智能的使用。

其次，必须制定相关法规和适当的立法，以保护隐私和数据安全。因此，必须确保所有相关行为体，包括提供数字技术的政府和公司，以道义和负责任的

方式利用人工智能，同时遵守《世界人权宣言》第十二条以及《公民及政治权利国际公约》第十七条概述的各项原则。

第三，应当推动全球数字协议，在先进国家与处于人工智能早期发展阶段的国家之间促进技术知识共享。人工智能专家、政府、公司和民间社会之间的协作努力旨在减轻滥用的风险，并促进负责任的人工智能实践。

总之，重要的是，要认识到人工智能所需的投入不能与现实世界脱节。数据、计算能力、电力、技能和技术基础设施等必要资源在全球的分布并不均衡。通过在人工智能优势与现有基本保障措施之间取得平衡，我们可以确保人工智能不会成为冲突的根源而加剧不平等和不对称，并对全球和平与安全构成潜在威胁。这种办法旨在利用人工智能的潜力，同时积极减轻可能出现的任何负面后果。

主席（*以英语发言*）：我现在请阿拉伯联合酋长国负责外交与国际先进科技合作部长助理发言。

沙拉夫先生（阿拉伯联合酋长国）（*以英语发言*）：首先，我要感谢古特雷斯秘书长今天所作专注议题的发言。主席先生，我感谢你和主席国联合王国在安理会讨论如此重要的议题。我也要感谢我们的其他通报者所作富有启发性的发言。

我们如何商讨应对人工智能的威胁和机遇，正迅速成为我们这个时代的决定性问题之一。五年前，阿拉伯联合酋长国和瑞士向古特雷斯秘书长提出建议，成立一个审议小组来审议这个问题。在秘书长的领导下，设立了数字合作高级别小组，从其审议中可以清楚地看出，人工智能等技术不能再不受检查了。

自计算机时代开始以来，计算处理能力一直遵循穆尔定律——每18个月翻一番。情况已不再如此。人工智能的发展正在超越穆尔定律，并以惊人的速度进展，各国政府都跟不上。这是我们需要的警钟。谈到人工智能，现在我们应当成为乐观的现实主义者，不仅是为了评估这项技术对国际和平与安全构成的威胁，也是为了利用这项技术所提供的各种机会。

为此，我今天将简要提出四点。

首先，我们必须确立道路规则。现在有一个短暂的机会之窗，让关键的利益攸关方愿意团结起来，考虑这项技术的防护栏。会员国应当从秘书长手上接过主导权，确立共同商定的规则来管理人工智能，以免为时太晚。这应当包括防止用人工智能工具宣扬仇恨、错误信息和虚假信息的机制，因为这些可能助长极端主义并加剧冲突。与其他网络技术一样，人工智能的使用应当受到国际法的严格指导，因为国际法继续适用于网络空间。但我们也必须认识到，可能必须采取一些策略，以便我们可以在人工智能快速发展的背景下有效适用国际法的常规原则。

其次，人工智能应当成为促进和平建设与冲突降级的工具，而不是加剧威胁的因素。人工智能驱动的工具具有更有效分析大量数据、趋势和模式的潜力。这可以提高实时发现恐怖活动以及预测气候变化如何对和平与安全产生不利影响的能力。它还为限制对攻击的错误归因铺平道路，并确保在冲突环境里采取相称的对策。与此同时，我们必须意识到，这项技术有可能被滥用来攻击关键基础设施并编造虚假故事，结果是，加剧紧张局势和煽动暴力。

第三，人工智能不应复制现实世界的偏见。如果我们不确保人工智能具有包容性，几十年来在消除歧视，特别是对妇女和女童的性别歧视以及对残疾人的歧视方面取得的进展将会遭到破坏。

数字合作高级别小组明确指出，包容性的数字经济和社会是必须立即予以关注的优先行动。人工智能提供的任何机会，只有设计和获得机会都以平等原则为基础，才能成为真正的机会。

第四，我们必须避免以阻碍创新的方式对人工智能进行过度监管。新兴国家在人工智能领域开展的创新、研究和开发活动对于这些国家的可持续增长和发展至关重要。为了保持这一趋势，新兴国家需要灵活性和敏捷的监管。我们应该培育一个鼓励负责任行为的部门，使用明智、有效和高效的法规和指南，避免可能阻碍这一技术发展的过于僵化的规则。

纵观历史，重大转变和飞跃往往在重大危机时刻后发生。第二次世界大战后联合国和安全理事会的创建就说明了这一事实。谈到人工智能，让我们不要坐等危机时刻发生。现在是走到危机前头，塑造一个着眼于维护国际和平与安全的人工智能舞台的时候了。

张军先生（中国）：中方欢迎阁下来安理会主持今天的会议。我感谢古特雷斯秘书长刚才所作的通报，他的很多看法建议值得深入研究。我也感谢曾毅教授、杰克·克拉克先生刚才所作的通报，他们的看法很有见地，有助于我们更好地理解和处理同人工智能有关的问题。

目前，人工智能迅猛发展、广泛应用，复杂效应不断显现。一方面，这一技术在科学研究、医疗护理、自动驾驶、智能决策等方面的赋能作用越来越突出，产生了巨大的技术红利。另一方面，人工智能应用范围在不断扩大，在侵犯数据隐私、散播虚假信息、加剧社会不平等、颠覆就业结构等方面引发越来越多的担忧。特别是人工智能一旦被误用，或被恐怖分子、极端势力滥用，将对国际和平与安全产生重大威胁。

目前，人工智能作为一门前沿技术，仍处在早期发展阶段。它作为“双刃剑”，究竟是好是坏、是善是恶，取决于人类如何利用、如何规范、如何统筹科学发展与安全。国际社会应该秉持真正的多边主义精神，开展广泛对话，不断凝聚共识，探讨制定人工智能治理的指导性原则。我们支持联合国在这方面发挥中心协调作用，支持古特雷斯秘书长召集各方进行深入讨论，支持各国特别是发展中国家充分参与、作出自己的贡献。

我谈几点初步看法。

第一，坚持伦理先行。人工智能的潜在影响可能超出人类的认知边界，要确保这一技术始终造福人类，就必须将“以人为本”、“智能向善”作为基本原则，用以规范人工智能的发展方向，避免这一技术成为“脱缰的野马”。在这两个准则的基础上，逐步建立并完善人工智能伦理规范、法律法规和政策体系，同时允许各国立足自身发展阶段和社会文化特点，建立符合自身国情的人工智能治理体系。

第二，坚持安全可控。人工智能相关技术发展和应用存在诸多的不确定性，安全是必须守住的底线。国际社会要强化风险意识，建立有效的风险预警和应对机制，确保不发生超出人类掌控的风险，确保不出现机器自主杀人。要加强对人工智能全生命周期的检测和评估，确保在关键时刻人类有能力摁下停止键。技术领先企业要明确责任主体，健全问责追责机制，避免发展或使用可能产生严重消极后果的风险技术。

第三，坚持公平普惠。人工智能在科技层面的影响是世界性、革命性的。发展中国家平等获取和利用人工智能技术、产品和服务，对于弥合南北技术鸿沟、数字鸿沟、发展鸿沟至关重要。国际社会要携手努力，确保发展中国家平等享受人工智能技术带来的发展红利，不断增强发展中国家在人工智能领域的代表性、发言权和决策权。个别发达国家为谋求科技霸权，构建排他性小圈子，以种种借口和行动恶意阻挠他国技术发展，人为制造科技壁垒，中方对此坚决反对。

第四，坚持包容开放。科技发展要在推动科技进步和保障安全应用两方面取得相对平衡，最佳路线是保持开放合作，鼓励跨学科、跨领域、跨地区、跨国界的交流对话，反对各种形式的“小院高墙”、“脱钩断链”。我们要在联合国框架下，推动国际组织、政府部门、科研和教育机构、企业、公众在人工智能发展与治理领域加强协调互动，共同打造开放、包容、公正、非歧视的科技发展环境。

第五，坚持和平利用。发展人工智能技术的根本目的是增进人类共同福祉，因此要重点挖掘人工智能在促进可持续发展、推动跨领域融合创新等方面的潜力，更好赋能全球发展事业。安理会可以深入研究人工智能在冲突局势中的应用和影响，采取措施丰富联合国的和平工具箱。人工智能在军事领域可能引发作战方式和战争形态的重大变革，各国应秉持负责任的国防政策，反对利用人工智能谋求军事霸权、破坏他国主权和领土完整，避免滥用、误用甚至恶用人工智能武器系统。

今天的会议关于人工智能问题的讨论更加凸显了构建人类命运共同体的重要性、必要性、紧迫性。中国秉持人类命运共同体理念，积极探索各领域人工智能发展与治理的科学路径。2017年，中国政府颁布《新一代人工智能发展规划》，明确提出科技引领、系统布局、市场主导、开源开放等基本原则。近年来，中国不断完善相关法律法规、伦理规范、知识产权标准、安全监测与评估措施等，充分保障人工智能健康有序发展。

中国始终以高度负责的态度参与人工智能全球合作与治理。早在2021年，中国在担任安理会轮值主席期间就主持召开了“新兴科技对国际和平与安全的影响”阿里亚模式会议，推动安理会首次聚焦人工智能等新兴科技问题。中国相继在联合国平台提交了关于人工智能军事应用、伦理治理的两份立场文件，从战略安全、军事政策、法律伦理、技术安全、规则制定、国际合作等角度提出系统的主张。

今年2月，中国政府发布了《全球安全倡议概念文件》，明确表示中方愿与国际社会就人工智能安全治理加强沟通交流，推动达成普遍参与的国际机制，形成具有广泛共识的治理框架和标准规范。我们愿同国际社会一道，在人工智能领域积极落实习近平主席提出的全球发展倡议、全球安全倡议、全球文明倡议，坚持发展优先，维护共同安全，促进跨文化交流与合作，携手各国共享人工智能技术的成果，共同防范和应对风险挑战。

德劳伦蒂斯先生（美利坚合众国）（*以英语发言*）：我感谢联合王国举行本次讨论，我也感谢秘书长、克拉克先生以及曾毅先生的宝贵见解。

人工智能给处理诸如粮食安全、教育以及医药待方面的全球性挑战带来了可喜的前景。自动化系统已经在帮助我们更高效地种植粮食，预测风暴路径，以及查明病人所患的疾病。因此，如果运用得当，人工智能可以加快我们实现可持续发展目标的进展。但是，它也有可能加剧威胁和冲突，包括通过散布虚假信息和错误信息，放大偏见与不平等，增加恶意的网络行动，以及加剧侵犯人权行为。因此，我们欢迎本次讨论，以期了解安理会如何能够找到人工智能裨益最大化与减少其风险之间的适当平衡。

安理会已经有处理两用能力、把变革性的技术融入我们维护国际和平与安全工作之中的经验。这些经验教导我们，成功来自于通过安全理事会和其它联合国机构，在正式和非正式场合与包括会员国、技术公司以及民间社会活动家在内的一系列行为体的合作。这正是美国承诺要做的，而且我们已经启动国内的努力。5月4日，拜登总统与主要的人工智能公司会面，强调我们各方都负有确保人工智能系统安全可信的重要责任。这些努力推进了美国国家标准与技术研究院的工作，最近该院发布了一个人工智能框架，为各种组织提供人工智能系统风险管理的自愿指导方针。通过白宫2022年10月份的一项人工智能权利法案蓝图，我们还确定了指导自动化系统的设计、使用以及部署的原则，以便涉及关键资源与服务的权利、机会以及准入能够得到平等享用和充分保护。现在，我们正与广泛的利益攸关方合作，以查明和处理有可能破坏和平与安全的涉及人工智能的人权风险。任何会员国不得利用人工智能来对民众进行审查、限制和压迫，或者剥夺民众的权能。

人工智能的军用可以而且应该做到合乎道义、负责任并且加强国际安全。今年早些时候，美国发布了一项关于人工智能和自动化的负责任军用的拟议政治宣言，阐述了如何根据适用的国际法在军事领域开发和使用人工智能的原则。该拟议宣言强调，人工智能能力的军用必须接受人类指挥链的问责，同时国家应采取步骤，最大限度地减少无意的偏见和事故。我们鼓励所有会员国认可该拟议宣言。

在联合国这里，我们欢迎努力开发和运用人工智能工具，以改进我们提供人道主义援助的共同努力，在气候变化和冲突等多样性问题上提供预警，并且推进其它共同目标。国际电信联盟最近的人工智能造福人类全球峰会是沿此方向迈出的一步。我们欢迎继续在安全理事会内部讨论如何管理技术进步，包括何时和如何采取行动，以处理政府或非国家行为体滥用人

工智能技术来破坏国际和平与安全的行为。我们还必须共同努力，以确保人工智能和其它新兴技术不被主要用作武器或者压迫的工具，而是被用作提升人类尊严、帮助我们实现理想、包括实现一个更加安全和平的世界的工具。美国期待与有关各方一道努力，确保负责任地开发和使用值得信赖的人工智能系统为全球造福。

弗兰萨·丹尼斯先生（巴西）（*以英语发言*）：我感谢秘书长今天通报情况和与会。我也感谢克拉克先生和曾毅先生的发言。

人工智能的迅速发展给加固我们的全球安全架构、拓宽决策进程以及加强人道主义努力带来了巨大潜力。我们还必须处理它带来的多层面的潜在挑战，包括在自主武器、网络威胁以及加剧现有不平等方面的挑战。随着我们着手讨论这个关键问题，我们要全面了解与人工智能有关的风险与机遇，努力利用其潜能为人类造福，同时确保维护和平、稳定以及人权。

我刚才宣读的这个段落完全是通过ChatGPT撰写的。尽管它存在概念上的不精准，但也显示出这些工具已经变得何其精密。技术的发展如此迅速，就连我们最优秀的研究人员也无法充分评估今后挑战的规模和这些技术可带来的裨益。今天任何讨论的起点必须基于虚心和这样一种认识，即：我们并不完全知道我们对人工智能缺乏哪些了解。我们确定知道的是，人工智能不是人类智能。大多数人工智能都依赖于大量数据，并通过复杂的算法建立模式和关系，使其能够生成上下文适当的结果。因此，结果在很大程度上取决于输入。人工监督对于避免偏差和错误至关重要。否则，我们将冒“垃圾进，垃圾出”的谚语成为自我实现的预言的风险。

与其它对安全有潜在影响的创新不同，人工智能主要是作为民用应用开发的。因此，主要从国际和平与安全的角度来看待它还为时过早，因为和平利用人工智能可能会对我们的社会产生最重大的影响。然而，我们可以肯定地预测，人工智能的应用将扩大到军事领域，对和平与安全产生相关后果。

虽然安理会应保持警惕，随时准备应对任何涉及使用人工智能的事件，但我们也应小心谨慎，不要通过在本会议厅集中讨论这一议题而使之过度安全化。人工智能涉及生活的方方面面，具有内在的多学科性质，因此，我们在国际上对人工智能的讨论应保持开放和包容。只有广泛而多种多样的观点才能让我们略知皮毛，并开始理解人工智能的不同层面。今天的通报会是一个良好的开端，将为我们带来有关人工智能发展和使用的各种观点。

尽管如此，鉴于人工智能的广泛影响和作用，成员组成具有普遍性的大会是最适合就人工智能进行有条理的长期讨论的论坛。目前正在进行的2021-2025年信息和通信技术安全和使用问题不限成员名额工作组将于下周举行第五届实质性会议，而人工智能是该工作组任务范围内各种主题中的一个重要议题。尽管地缘政治环境充满挑战，但向所有会员国开放的该工作组在逐步就与国际和平与安全有关的信息和通信技术问题达成全球共识方面取得了进展。考虑到网络技术的特殊性，这正是我们在讨论网络技术带来的挑战时所应追求的目标。

人工智能的军事应用，特别是在使用武力方面，应严格遵守《日内瓦四公约》和其它相关国际承诺所载的国际人道法。巴西一贯以“有意义的人类控制概念”为指导。《特定常规武器公约》缔约方2019年批准的指导原则（b）指出：“必须保留人类对决定使用武器系统的责任，因为不能把责任转嫁给机器”（CCW/MSP/2019/9，附件三）。

人的因素在任何自主系统中的中心地位对于建立道德标准和充分遵守国际人道法至关重要。人的判断和问责是不可替代的。

人工智能的军事应用必须基于其从开发到部署和使用的整个生命周期的透明度和问责。此外，具有自主功能的武器系统应消除系统操作中的偏差。我们必须通过强有力的规范来防止偏差和滥用，并保证遵守国际法，特别是国际人道法和人权法，从而迅速推进逐步制定有关自主武器系统使用的条例和规范的

工作。在各国使用人工智能技术时，以及在安理会希望在其维和任务或维护国际和平与安全的更广泛任务中使用这些技术时，必须遵守国际法。

除了具有自主功能的常规武器带来的挑战外，我们不应该回避就人工智能和大规模毁灭性武器的相互作用带来的内在风险发出非常严厉的警告。我们震惊地注意到，有消息称，人工智能辅助的计算机系统能够在几个小时内开发出新的有毒化合物，并设计出新的病原体和分子。我们也不应该让核武器与人工智能联系在一起的可能性危及我们共同的未来。

人工智能具有在未来几年既重塑又破坏我们的社会的巨大潜力。要在这两者之间取得平衡，需要国际社会做出广泛一致的努力，这将包括但绝不仅限于安全理事会。联合国仍然是唯一能够促进监督和塑造人工智能发展，以便人工智能能够造福人类，并符合使我们聚集在这里的共同宗旨和原则所需的全球协调的组织。

贝里斯维尔夫人（瑞士）（*以法语发言*）：我们感谢安东尼奥·古特雷斯秘书长参加本次重要辩论。我还要感谢克拉克先生和曾毅先生作了宝贵和令人印象深刻的发言。

（*以英语发言*）

“我相信，我们看到成千上万个像我一样的机器人发挥作用，只是时间问题”。

（*以法语发言*）

这些话是机器人Ameca在秘书长提到的、由国际电信联盟和瑞士两周前在日内瓦共同举办的人工智能造福人类全球峰会上对一名记者说的。

虽然人工智能因其速度快且显然无所不知而可能是一个挑战，但它也能够而且必须为和平与安全服务。当我们展望《新和平纲领》时，我们掌握人工智能，以确保其造福人类而不是损害人类。在这种背景下，让我们抓住机会，通过与前沿研究密切合作，为人工智能造福人类铺平道路。为此，位于苏黎世的瑞士联邦理工学院正在为联合国危机和行动中心开发一个原型人工智能辅助分析工具。该工具将探索人工智能在维和，特别是在保护平民和维和人员方面的潜力。此外，瑞士最近发起了一项关于信任和透明倡议的呼吁，学术界、私营部门和外交部门可在该倡议中共同寻求切实可行的快速解决方案，以应对与人工智能相关的风险。

安理会也必须努力应对人工智能给和平带来的风险。因此，我们非常感谢联合王国组织本次重要辩论会。让我们以网络操作和虚假信息为例。虚假叙述破坏了人们对政府和和平特派团的信任。在这方面，人工智能是一把双刃剑。虽然它突出虚假信息，但也可以用来检测虚假叙述和仇恨言论。那么，我们如何才能收获人工智能为和平与安全带来的好处，同时将风险降至最低呢？我想提出三条途径。

第一，我们需要一个由参与开发和应用这种技术的所有行为体——政府、企业、民间社会和研究组织——共享的共同框架。我认为，克拉克先生在他的通报中非常清楚地表明了这一点。人工智能并非存在于规范真空中。包括《联合国宪章》、国际人道法和人权在内的现有国际法都适用于它。瑞士致力于所有有助于重申和澄清人工智能国际法律框架的联合国进程，并在涉及致命自主武器系统时，致力于制定禁止和限制措施。

第二，人工智能必须以人为中心，或者如曾先生先前所说，

（*以英语发言*）

“人工智能永远不应该假装是人”。

（*以法语发言*）

我们呼吁人工智能的开发、部署和使用始终以伦理和包容性考虑为指导。必须保持对政府、公司和个人的明确责任和问责。

最后，人工智能发展处于相对早期阶段，这让我们有机会确保平等和包容，反击歧视性成见。我们提供的数据优质可靠，人工智能才会优质可靠。如果数据体现偏见和成见，例如性别偏见和成见，或者说，

如果它根本不代表其操作环境，人工智能就会向我们提供关于维护和平与安全的错误建议。政府和非政府开发者及用户都有责任确保人工智能不会复制我们正在努力克服的有害社会偏见。

安全理事会有责任主动监测人工智能的有关动态以及它可能对维护国际和平与安全构成的威胁。它应当以大会在有关法律框架方面取得的成果为指导。安理会还必须运用其权力，通过预测风险和机遇，或者鼓励秘书处及和平特派团以创造性和负责任的方式利用这种技术，确保人工智能为和平服务。

我国代表团在利用人工智能召开了我国担任主席国期间的首次辩论会，探讨为保持和平建立面向未来的信任问题（见S/PV.9315），也利用人工智能同红十字国际委员会合作举办了关于数字困境问题的展览。我们看到了这种技术促进和平的非凡潜力。因此，我们期待“智能向善”成为《新和平纲领》不可或缺的内容。

阿杰曼先生（加纳）（*以英语发言*）：我感谢联合王国召开本次关于人工智能的高级别辩论会，感谢秘书长安东尼奥·古特雷斯今天上午向安理会所作的重要通报。我们同样感谢杰克·克拉克先生和曾毅先生提供专家意见。

人工智能是我们社会中无处不在的结构物，它正在显现的支配性优势可能对若干领域产生积极影响，包括在医药、农业、环境管理、研究开发、文化艺术领域以及贸易方面的有益应用。我们看到机遇在望，那就是可以通过进一步应用人工智能来加强各方面生活的成果，同时我们也可能已经看到一些危险，为此，我们大家必须行动起来，迅速采取合作方式来努力避免可能危害我们全人类的风险。

人工智能——特别是在和平与安全方面利用人工智能——的指导原则必须是，不能复制一些强大技术因为能够引发全球规模的灾难而给全世界造成的风险。我们必须约束个别国家谋求作战优势的过度野心，并承诺努力制订为出于和平目的管理人工智能技术的原则和框架。

对加纳而言，我们看到有机会发展和应用人工智能技术来识别冲突预警信号，并制订成功率更高、而且可能更具有成本效益的应对措施。这类技术可能有助于协调人道主义援助并改进风险评估。它们在执法方面的应用已经在很多司法辖区内很受赞赏，那里的执法卓有成效，冲突风险通常较低。

此外，运用人工智能技术开展和平调解和谈判工作，取得了为了和平事业所必须实现的令人瞩目的初步成果。例如，部署人工智能技术来确定利比亚民众对政策的反应，已经促进了和平，该国的2022年全球和平指数排名上升，反映了这一点。我们也认为，可以在相似的情况下，在维和特派团的情报、监视和侦查职能中，通过负责任地配置人工智能技术来提高维和人员的安全保障和对平民的保护。

尽管这些动态令人鼓舞，我们仍然认为国家行为体和非国家行为体使用人工智能会带来风险。人工智能技术与自主武器系统相结合，是令人关切的首要问题。虽然寻求发展此类武器系统的国家的真正意图可能是减少它们卷入冲突所产生的人员伤亡，但是这戳穿了它们对和平世界的虚假承诺。我们所了解的人类掌握操控原子能力的历史显示，如果这种欲望继续存在，它们只会促使其他国家同样努力消除这类威慑力量所寻求创造的优势。非人类控制这些武器系统产生的额外危险也是一种全世界承受不起也无法忽视的风险。

日益数字化的世界和创造虚拟现实意味着分辨真实与虚构的能力日益削弱。这可以创造不受制约的平台，尤其是非国家行为体可以利用人工智能技术把这些平台当作工具，以破坏社会稳定或引起国家之间的摩擦。虽然人工智能技术可用于打击错误信息、虚假信息和仇恨言论，但是它们也有可能被恶势力用来实现这些势力的邪恶图谋。

然而，智能向善的潜力应当引导我们努力实现其和平用途。正如早先指出的那样，必须制订一些原则和框架，同时考虑到我们尚未充分了解人工智能技术的演变。尽管如此，这样一个进程不应当只由安全理

事会来处理，而应当由联合国广大会员国处理，因为我们如何引导人工智能技术的进一步演变，同样与广大会员国利害相关。没有全球共识，很难限制人工智能技术的蓬勃发展。

目前，人工智能技术的发展有很大一部分是在私营部门和学术界发生的，因此还必须将对话扩展到各国政府以外，以确保在弥补产业差距的同时，不转用或滥用包括无人驾驶飞行器在内的人工智能技术。应当预料到并减轻这些飞行器对和平与安全的消极影响，包括在非洲大陆的消极影响，那里的恐怖团体可能正在实验这类技术。

认识到人工智能技术可能打破军事平衡，各国必须有计划地采取建立信任措施，这些措施基于一种共同意愿，即防止非故意引发的冲突。通过制订关于各国实施的借助人工智能的系统、战略、政策和方案的自愿信息分享和通知标准，能够做到这一点。我们希望会员国在审议秘书长即将发表的关于《新和平纲领》的政策简报时，能够推动持久解决方案，以应对国际和平与安全面临的新威胁。我们希望对秘书长在这方面所作的努力表示支持。

在此过程中，我们还必须深入开展工作，推动现有倡议和当前进程，例如秘书长数字合作路线图，当前关于缔结一项打击为犯罪目的使用信息和通信技术行为全球公约的谈判，以及通信技术安全和使用问题不限成员名额工作组。同样，我们鼓励安全理事会在“以行动促维和+”倡议之下进一步参与联合国维持和平行动数字化转型战略。在即将于阿克拉召开联合国维持和平部长级会议期间，加纳将欢迎举行对话，讨论在相关主题之下如何部署人工智能来加强维和行动。在非洲，非洲联盟数字转型战略（2022-2030年）也将继续成为非洲大陆自由贸易区的一项重要附属战略，而非洲大陆自由贸易区是应对非洲大陆诸多深层安全挑战和平息非洲枪炮声的支撑点。

最后，我申明加纳致力于推动关于利用人工智能技术促进我们的世界和平与安全的建设性讨论。我们强调，必须采取全社会联动的办法，发挥私营部门特别是技术巨头的潜力，并继续将公民的人权置于所有道德原则的核心。

德里维埃先生（法国）（*以法语发言*）：我谨感谢秘书长、杰克·克拉克先生和曾毅先生的通报。

人工智能是二十一世纪的革命。当我们面临一个以竞争和混合战争为特征的更加严峻的世界时，必须使人工智能成为服务于和平的工具。

法国相信，人工智能可以在维护和平方面发挥决定性作用。这些技术可以促进维和人员的安全和行动绩效，特别是改善对平民的保护。它们也可以通过动员民间社会——未来也许可以通过促进提供人道主义援助——来帮助解决冲突。

人工智能也可以为可持续发展目标服务。这是我们对秘书长《全球数字契约》的贡献的主旨。在应对气候变化的斗争中，人工智能可以通过提供更准确的天气预报来帮助我们预防自然灾害。它也可以支持履行在《巴黎协定》中做出的减少温室气体的承诺。

人工智能的发展也带来风险，我们必须直面这些风险。人工智能可能会增加网络威胁，因为它使恶意行为体能够实施越来越复杂的网络攻击。人工智能系统本身也容易受到网络攻击。因此，确保其安全至关重要。

在军事领域，人工智能可以深刻改变冲突的性质。因此，我们必须在《特定常规武器公约》致命自主武器系统政府专家组内继续努力，制定适用于此类系统的框架。这一框架必须确保未来的冲突遵守国际人道法。

最重要的是，生成式人工智能可以通过低成本大规模传播人工内容或根据受众调制的信息来加强信息战。我们只需看看中非共和国和马里正在进行的大规模造谣活动，或者俄罗斯对乌克兰发动战争的同时开展的造谣活动，就会知道这一点。外国干涉选举的活动破坏国家稳定，并使民主的基础受到质疑。

法国致力于促进以合乎道德和负责任的方式对待人工智能。这是我们在2020年发起的全球伙伴关系

的目标。在欧洲联盟和欧洲委员会内，法国正在制定管理和支持人工智能发展的规则。

面对这场革命，联合国提供了一个不可替代的框架。我们欢迎《新和平纲领》正在进行的工作和即将举行的未来峰会，该会议将使我们能够集体思考这些问题并制定未来的标准。法国将尽最大努力让人工智能为预防冲突、维持和平和建设和平服务。

佩雷斯·卢塞先生（厄瓜多尔）（*以西班牙语发言*）：波兰作家斯坦尼斯拉夫·莱姆说过：“智能的首要义务是自我怀疑”。但我们不能指望人工智能做到这一点。

因此，我强调，在联合王国倡议下召开的本次会议的主题具有现实意义，我也感谢安东尼奥·古特雷斯秘书长和其他发言者所作的通报。

问题不能是我们是否支持人工智能的发展。在技术快速变化的背景下，人工智能已经在以惊人的速度发展，并将继续飞速发展。

同其他任何技术一样，人工智能是一种工具，可以促进维持和平和建设和平的努力，也可以被用来破坏这些目标。人工智能可以有助于预防冲突和在诸如2019冠状病毒病等复杂背景下主持对话。新兴技术对于克服疫情构成的障碍至关重要。

人工智能可以支持保护人道主义人员，使准入和行动得以扩大，包括通过预测分析实现这一目标。备灾、预警和及时反应可以受益于这一工具。技术解决方案可以帮助联合国维和行动更加有效地完成任务，包括促进适应不断变化的冲突动态。

2020年3月30日，厄瓜多尔联署了第2518（2020）号决议，该决议支持更加综合地利用新技术，以加强人员对态势的感知并提高其能力，2021年5月24日的主席声明（S/PRST/2021/11）重申了这一点。这必须包括人工智能，因为它有能力监测和分析冲突，改善营地和车队的安全。

作为一个组织，我们如果没有克服新的安全挑战的工具，就无法提高效率。我们的责任是促进和利用技术发展，将其作为和平的促进因素。在此过程中必须严格遵守国际公法、国际人权法和国际人道法。我们不能忽视出于恶意或恐怖主义目的误用或滥用人工智能所带来的威胁。人工智能系统也带来其他风险，比如歧视或大规模监控。

此外，厄瓜多尔反对将人工智能军事化或放置人工智能武器。我们重申致命自主武器系统带来的风险，并重申所有武器系统都必须在唯一可行的责任和问责框架下，服从人类的决策、控制和判断。

道德原则和负责任的行为不可或缺，但还不够。正如厄瓜多尔将继续倡导的那样，要利用人工智能而不加剧人工智能产生的威胁，答案是建立一个具有法律约束力的国际框架。此外，在无法确保人类充分控制致命自主武器以及无法确保区分、相称和预防原则的情况下，应该禁止此类武器。

我同秘书长一样感到关切的是，人工智能与核武器之间可能存在令人震惊的联系。我们欢迎今天就《新和平纲领》提出的建议，并同意，必须弥合数字鸿沟，促进伙伴关系和联系，使我们能够利用现有技术实现和平目的。除道德考虑之外，冲突的机器人化对裁军努力构成重大挑战，也构成生存挑战，这一点不容安理会忽视。

今天的人工智能研究人员谈到对齐的问题，即我们如何确保人工智能的发现为我们服务，而不是摧毁我们？这是爱因斯坦、奥本海默和冯·诺依曼等科学家继而扪心自问的问题。是我们今天面临的挑战。

弗雷泽夫人（马耳他）（*以英语发言*）：我感谢主席国联合王国今天就这个非常热门的问题举行通报会。我也感谢秘书长以其想法和见解丰富了我们的讨论。

人工智能正在重塑我们的工作、互动和生活方式。人工智能的和平应用有助于实现可持续发展目标，并支持联合国的维和努力。这些努力包括使用无人机运送人道主义援助、进行监测和监视。

另一方面，人工智能技术的扩散带来了需要我们

关注的重大风险。人工智能的潜在滥用或意外后果，如果不加以谨慎管理，可能对国际和平与安全构成威胁。恶意行为体可能利用人工智能进行网络攻击、宣传虚假信息和错误信息或将其用于自主武器系统，导致脆弱性和地缘政治紧张局势增加。人工智能还可能对人权造成负面影响，包括歧视性的算法决策所带来的后果。我们必须通过国际合作、框架和规范来集体应对这些风险。

马耳他认为，国际、区域和国家各级和各部门的多个利益攸关方的合作对于在全世界实施人工智能伦理框架至关重要。在这方面，国际社会必须制定全球文书，这些文书不仅要注重阐述价值观和原则，还要注重这些价值观和原则的实际实现，并大力强调法治、人权、性别平等和环境保护。在政府和非政府行为体争先发展人工智能的同时，治理和控制做法也必须以同样的速度发展，以维护国际和平与安全。因此，安全理事会必须推动强有力的人工智能治理，并通过分享经验和政府框架，确保以包容、安全和负责任的方式部署人工智能。

自2019年以来，马耳他一直在根据欧洲联盟（欧盟）《可信赖人工智能伦理准则》制定人工智能伦理框架。该框架基于四项指导原则：第一，在以人为本的基础上建立人工智能；第二，尊重所有适用的法律法规、人权和民主价值观；第三，最大限度地发挥人工智能系统的效益，同时预防并尽量降低其风险；第四，与人工智能伦理相关的新兴国际标准和规范保持一致。马耳他愿意在人工智能问题上与所有相关利益攸关方携手合作，就负责任地使用人工智能的共同标准达成全球协议。此外，在欧盟内部，我们正在制定一项人工智能法案，旨在确保公民能够信任人工智能所能提供的服务。该法案以基本权利和法治为基础，对人工智能采取以人为本、有利于创新的做法。

在这方面，马耳他大力支持2021-2025年信息和通信技术安全和使用问题不限成员名额工作组的工作，并强调建立信任措施对于提高对话水平和信任程度至关重要，以便在使用人工智能方面提高透明度，确保更好地问责。

马耳他对在军事行动中使用人工智能系统的行为表示关切，因为机器无法像人一样做出涉及区别、相称和预防等法律原则的决定。我们认为，目前利用人工智能的致命自主武器系统应予以禁止，只有那些完全遵守国际人道法和人权法的武器系统才应受到管制。同样，将人工智能纳入国家安全、反恐和执法系统的做法引发了基本人权、透明度和隐私方面的问题，这些问题必须得到解决。

最后，马耳他认为，安全理事会在这个问题上可以发挥重要的前瞻性作用。我们有责任密切监测事态发展，并及时应对可能出现的对国际和平与安全的任何威胁。我们只有通过促进负责任的治理、国际合作和道德考虑，才能在减轻潜在风险的同时利用人工智能的变革力量。

恩格耶马·恩东夫人（加蓬）（*以法语发言*）：我谨感谢联合王国在技术创新不断增加、革新我们的社会并对国际安全产生影响之际，组织本次关于人工智能的辩论会。我也感谢安东尼奥·古特雷斯秘书长、曾毅教授和杰克·克拉克先生所作的通报。

人工智能既迷人又令人困惑。近年来，它彻底改变了我们的生活方式、生产方式和思维方式，并扩大了我们现实的边界。由于其精确性和解决复杂问题的能力，人工智能系统从更传统的信息技术机制中脱颖而出，为维护国际和平与安全提供了许多机会。

多年来，维护国际和平与安全一直依赖于强大的技术生态系统，该系统不仅能提高危机管理和预防能力，还能促进增强对实地局势的了解，同时改善对平民的保护，特别是在复杂环境中的保护。

人工智能正在通过成倍提高预警系统的分析能力来作出自己的具体贡献。现在，通过在极短的时间内分析各种来源的大量数据，可以更快、更容易地发现新出现的威胁。得益于人工智能，联合国和平特派团的业务系统正变得越来越有效。的确，例如使用无人机、夜视和地理定位系统，就有可能探测武装团体和恐怖团体的活动，确保在难以到达的地区运送人道主义援助，并改善停火监测和实地探雷任务。它还加

强了非常复杂的维和任务, 特别是保护平民的任务的执行力度。

人工智能在建设和平进程中也可以发挥重要作用, 有助于冲突后国家的重建工作, 并促进实施速效项目, 同时为年轻人提供就业机会, 为前战斗人员提供重返社会的可能性。然而, 为了最大限度地发挥人工智能对和平与安全的益处, 特别是在部署维持和平行动时, 当地社区必须适用和吸收这些新技术, 以便在国际部队撤出后保持其有益效果。如果这些技术未能在地方一级扎根, 那么人工智能的益处很可能会消失, 危机也会重新出现。各国、国家和国际组织以及当地居民都必须在制造与分销过程中接受教育, 以加强对所使用的人工智能系统的信任以及这些系统的合法性。

人工智能当然有助于加强国际和平与安全, 但它也带来了许多风险, 我们现在必须了解这些风险。恐怖主义和犯罪集团可以利用人工智能提供的许多机会从事非法活动。近年来, 黑客网络加大了网络攻击、造谣和窃取敏感数据的力度。恶意使用人工智能带来的威胁应该为国际社会敲响警钟, 并作为加强对新技术发展控制的出发点。这意味着以联合国为保证方, 加强透明度和国际治理, 但最重要的是加强问责。联合国必须加强国际合作, 制定一个具有适当控制机制和强有力安全系统的监管框架。共享信息和订立道德标准也将有助于防止滥用和维护国际和平与安全。

加蓬仍然致力于促进和平和负责任地使用新技术, 包括人工智能。有鉴于此, 重要的是要鼓励分享安全和控制领域的最佳做法, 鼓励各国采取国家监管政策, 以及立即启动提高认识方案, 特别是提高年轻人对人工智能问题和挑战的认识。

总之, 人工智能显然提供了一系列机遇。它支持可持续发展倡议, 有助于预防人道主义和安全危机以及应对气候变化及其负面影响。然而, 如果缺乏可靠的条例和有效的控制与管理工具, 人工智能可能会对国际和平与安全构成真正的威胁。因此, 对我们对这些日益尖端的技术的热情必须保持谨慎和克制。

霍查先生(阿尔巴尼亚)(*以英语发言*): 主席女士, 我感谢你召开今天的重要会议并将这个问题提交给安全理事会, 以便就该问题进行首次辩论。

作为世界科技驱动力的一部分, 人工智能已经存在了几十年。正如秘书长、我们的通报人和其他同事所说的那样, 人工智能的发展最近突飞猛进, 为其在人类活动的几乎每一个部门的使用开辟了广阔的途径。一切都表明, 在未来几年, 它将在一个革命性的、前所未有的技术进步时代占据前沿和中心位置。我们的世界对科学进化和颠覆性技术增长并不陌生。它是人类不间断地追求进步的一部分, 是我们基因组的一部分, 并铭刻在人类历史上。但是, 人工智能有着本质上的不同。它在进展速度和潜在应用范围方面都非常突出, 并有望前所未有地改变世界, 并以我们现在甚至无法想象的规模实现流程自动化。

当这项技术以令人惊叹的速度发展时, 我们陷入迷恋与恐惧之中, 权衡利益与担忧, 期待可能改变世界的应用, 但也意识到它的另一面——阴暗面, 即可能对我们的安全、隐私、经济和安保产生影响的潜在风险。有些人更加危言耸听, 甚至警告人工智能可能给我们的文明带来风险。一项带来我们无法完全掌握的影响深远的后果的技术以令人眼花缭乱的节奏发展, 这正确地引发了严重的问题。这是因为技术性质, 算法如何得出结果缺乏透明度和问责, 以及甚至连设计人工智能模型的科学家和工程师也常常不完全理解它们是如何得出其生成的输出结果的。

视对其进行训练所用的数据集或组织其算法的方式而定, 人工智能模型和系统可能会导致基于种族、性别、年龄或残疾的歧视。虽然这些问题是可以预防或纠正的, 但要预防那些使用这项技术意图造成伤害的人所带来的严重风险则要困难得多。互联网和社交媒体已经显示此类行为的危害性有多大。一些国家不断企图通过滥用数字技术来蓄意制造误导性信息、歪曲事实、干涉他国民主进程、散布仇恨、宣扬歧视、煽动暴力或冲突。深度伪造和篡改照片正被用来制造令人信服但虚假的信息和叙事,制造令人信服的阴谋论, 破坏公众信任和民主, 甚至引发恐慌。对于所有这些

行为体来说，人工智能将为恶意活动提供无限可能。

滥用人工智能可能对国际和平与安全产生直接影响，并构成严重的安全挑战，而我们目前对此准备不足。人工智能可能被用于通过大规模虚假信息攻击来延续偏见、开发新的网络武器、为自主武器提供动力以及设计先进的生物武器。

当我们收获技术进步的好处时，当务之急是利用现有的规则和条例，改进和更新它们，并确定使用人工智能的道德规范。我们还必须在国家和国际一级建立必要的保障措施、治理框架以及明确的责任和权限分工，以确保适当、安全和负责任地使用人工智能系统，造福于所有人，并确保人工智能系统不侵犯人权和自由或影响和平与安全。我们必须促进负责任的国家行为标准和国际法在使用人工智能及其技术方面以及在监测和评估风险和影响方面的适用性。这为安理会提供了一个作用。阿尔巴尼亚将继续促进人权、基本自由和法治在其中得到尊重的开放、自由和安全的人工智能技术。

波利扬斯基先生（俄罗斯联邦）（*以俄语发言*）：我们欢迎秘书长参加今天的会议，并认真听取了他的发言。我们也感谢各位通报人提出了令人感兴趣的意见。

俄罗斯联邦高度重视发展旨在为人类共同利益服务和促进社会经济进步的先进技术。人工智能是一项开创性的现代技术，无疑激发了相当大的科学兴趣。然而，它的实际意义和潜在应用，更不用说它对政治进程的假设影响，还有待全面评估。

我们从今天的通报人那里听到的评估给我们留下的印象是，作为一项技术，人工智能正处于发展的早期阶段。未知数仍然远远多于已知数。因此，选择安全理事会会议厅作为今天讨论的平台，对我们来说并不明显。即使在今天会议的概念说明（S/2023/528）中，显然也没有具体的论据来支持人工智能与国际和平与安全问题之间有机联系的概念，至少现在还没有。概念说明只是列举了一组可以应用于人们想要应用的任何领域的不同技术趋势。

在准备今天的会议时，我决定询问主要信息来源——人工智能本身——人工智能是否对国际和平与安全构成威胁。这是它的回答。

> “人工智能可能会对国际和平与安全构成某些威胁，但它也可以产生积极影响。以下是应该考虑的几点。
>
> “一是网络安全。人工智能发展可能会在网络空间造成新的漏洞和威胁。人工智能可被用来制造强大的网络武器，实现网络攻击自动化，以及规避现有的防御系统。
>
> “二是自主武器系统。开发基于人工智能的自主武器系统可能会带来安全风险，因为这类系统可以在没有人类直接参与的情况下决定使用武力。这可能导致不可预测的后果，并加剧武装冲突。
>
> “第三也是最后一点是虚假信息的传播。人工智能可被用来制造和传播虚假信息和假新闻，这可能会破坏信息的可信度，并加剧社会和政治紧张局势”。

因此，我们认为，试图在人工智能与对和平与安全的威胁之间建立普遍联系的做法有点做作。我们可以看到，人工智能同意我们的观点。

首先，国际社会应确定可能存在的风险和威胁的性质和特点，评估其规模和可能的应对方式等。这需要在科学专门知识的基础上进行专业讨论，这可能需要数年时间。这项工作已经在专业平台上进行。也正在以专门形式对可能对全球和区域安全产生影响的人工智能的各种军事层面进行讨论。致命自主武器系统问题属于致命自主武器系统问题政府专家组和《特定常规武器公约》（又称《不人道武器公约》）缔约国的职权范围。大会主持的2021-2025年信息和通信技术安全和使用问题不限成员名额工作组全面讨论了与使用信息和通信技术有关的安全问题。我们认为，重复他们的工作将适得其反。

与任何形式的先进技术一样，人工智能可能对人

类产生有益的后果，也可能产生破坏性的后果，这取决于谁控制它以及用于何种用途。今天，我们不幸地看到，以美国为首的西方正在破坏对其自身的技术决策和执行这些决策的信息技术公司的信任。我们经常看到有证据表明美国特勤部门干预该行业主要厂商的活动，操纵内容审核算法并跟踪用户，包括通过制造商的内置硬件和软件功能来这样做。然而，西方认为，如果人工智能遵循适合自己的政治议程，那么，故意允许人工智能忽略社交网络平台上的仇恨言论，就不会存在道德问题，就像极端主义公司 Meta 及其允许发布杀害死俄罗斯人的呼吁的情况一样。与此同时，这种算法学会了发布虚假和错误信息，并自动屏蔽社交网络所有者及其在情报机构的操纵者认为错误的信息--换句话说，令人痛苦的真相。本着臭名昭著的取消文化的精神，人工智能被要求按需编辑整个数字阵列，从而产生虚假的叙述。简而言之，挑战和威胁的主要来源并不是人工智能本身，而是所谓的先进民主国家中人工智能的肆无忌惮的倡导者。讨论这一点与讨论主席国联合王国作为召开本次会议的理由所列举的问题同样重要。

如今，人们普遍认为，人工智能将为新兴市场和财富来源创造巨大前景。然而，这种潜在利益的不平等分配问题被小心翼翼地回避了。秘书长在其《数字合作路线图》报告中详细讨论了这些方面。数字鸿沟已经达到这样一个程度，在欧洲，89%的人可以上网，而在低收入国家，这一数字只有25%。全球近三分之二的商业和服务是在数字平台上进行的，然而在南亚和撒哈拉以南非洲，一部智能手机的费用占平均月收入的40%以上，同时，非洲用户的移动数据费用是全球平均水平的三倍多。最后，世界上只有不到一半的国家提供了旨在使公民掌握数字技能的政府支持。

这是因为创新创造的财富分配极为不均，由少数几个主要平台和国家掌控。数字技术显著提高了生产率和附加值，但其好处并未带来共同繁荣。联合国贸易和发展会议的《2023年技术和创新报告》警告称，发达国家将从包括人工智能在内的数字技术中受益最大。数字技术正在帮助将经济权力集中在越来越少的精英和公司手中。2022年，科技亿万富翁的总财富为2.1万亿美元。这种差异的背后是治理和公共投资方面的巨大差距，特别是在跨境治理方面的巨大差距。

从历史上看，数字技术一直是由私营部门开发的，政府在出于公共利益对其进行监管方面一直落在后面。这一趋势必须扭转。国家应在制定人工智能监管机制方面发挥领导作用。该行业采用的任何自律工具都应符合这些公司所在国家的国家立法。我们反对建立人工智能的超国家监督机构。我们还认为，在这一领域域外强加规范是不可接受的。只有在国际社会主权成员之间相互尊重、平等对话的基础上，并适当考虑谈判进程参与者的所有合法利益和关切，才有可能在这一领域达成普遍协议。俄罗斯已经在为这一进程作出贡献。在我国，各大信息技术公司在人工智能领域制定了国家道德准则，为人工智能系统的安全、合乎道德的开发和使用制定了指导方针。它不规定任何法律义务，对外国专门机构、私营公司以及学术和社会实体开放。该准则是国家努力制定的，旨在促进教科文组织《人工智能伦理问题建议书》的实施。

最后，我想强调的是，不应允许任何人工智能系统质疑人类的道德和智力自主性。开发人员应定期评估与使用人工智能相关的风险，并采取措施将风险降至最低。

主席（以英语发言）：发言名单上没有其他发言者了。

我再次感谢我们的技术专家与会，感谢我们的同事们今天所作的发言。

中午12时15分散会。