中方关于联合国“军事人工智能对国际和平与安全的影响”非正式会议的意见

人工智能引领新一轮科技和产业革命，重塑战争形态和国防建设范式，也带来新的风险挑战。

中方认为，

在军事人工智能全球治理这一涉及到人类前途命运的重大问题上，联合国应当发挥应有作用。本次非正式会议契合全球应对风险挑战的现实需求，为各成员国搭建平等对话平台，中方对此予以肯定。同时，为提升军事人工智能全球治理的务实成效，应做好与“致命性自主武器系统”政府专家组等成熟机制对接协同，避免资源分散，防止治理碎片化。

在当前背景下，各成员国应秉持共同、综合、合作、可持续的全球安全观，通过多边协商，就如何规范人工智能军事应用寻求共识，构建有效的治理框架，维护全球战略稳定、防止军备竞赛、缓解人道主义关切，在人工智能领域践行构建人类命运共同体。国家应该完全承担人工智能军事应用的治理责任。多利益攸关方见解可作为参考，但同时必须防止商业导向左右治理规则设计。

中方认为，

应当客观评价军事人工智能的意义，既要引导其正确发展方向，也要防范其毫无约束的发展。中国国家主席习近平提出全球治理倡议，从奉行主权平等、遵守国际法治、践行多边主义、倡导以人为本、注重行动导向五个方面阐释重要内涵，为包括人工智能在内的新兴领域全球治理提供了根本遵循。

各方在治理路径仍存分歧的情况下，中方认为可先确立指导原则，在此基础上不断凝聚共识，完善具体治理举措。秉持遵守联合国宪章宗旨和原则、恪守国际关系基本准则理念，结合此前中方提出的《全球人工智能治理倡议》《关于规范人工智能军事应用的立场文件》，中方在此次会议上提出坚持“和平正义”“慎重负责”“人道主义”“多边协商”等“四个坚持”，欢迎各方多提宝贵意见。

中方注意到，

诸多代表团就个别国家借助军事人工智能谋求绝对优势，纵容人工智能巨头跨域介入战争、侵犯他国主权、造成人道主义灾难等新动向表达了关切，这反映了国际社会的正义呼声。

中方认为，各方应高度关注军事人工智能对全球战略平衡和稳定、国际和平与安全带来的风险挑战，后续可就此深入研究探讨。各国特别是大国应确保人工智能军事应用安全

可靠可控，不应将之作为挑起地区矛盾、破坏国际和地区和平稳定、以大欺小、倚强凌弱的工具，更不能将战场作为试验场，这是“负责任”的核心要义。

中方注意到，

本次会议上，各方普遍关注某些发达国家垄断技术和算法资源，扩大人工智能技术鸿沟，损害发展中国家和平利用科技权利。

中方一贯认为，考虑到人工智能技术的军民两用性质，在加强监管和治理的同时，避免采取以意识形态划线、泛化国家安全概念的做法，消除人为制造的科技壁垒，确保各国充分享有技术发展与和平利用的权利。各国应坚持开放包容原则，开展对话交流，增进相互理解，在治理问题上各国应加强政策协调和能力建设合作，不断提升治理水平。中方愿积极支持联合国开展军事人工智能国际合作与能力建设相关工作，为推动建立普遍参与、广泛共识的治理规范贡献力量。

中国代表团在联合国“军事人工智能对国际和平与安全的影响”非正式会议上的发言

Session 1：Opening Remarks and Scene Setting

感谢联合国裁军办组织此次会议。

军事智能化已成为当今全球军事变革大势。人工智能引领新一轮科技和产业革命，重塑战争形态和国防建设范式，也带来新的风险挑战。

个别国家囿于维霸执念，纵容科技巨头介入战争、成为战争机器；任由算法越境操纵，侵犯他国主权；虚化人类控制底线，酿成人道主义悲剧。完善军事人工智能全球治理时不我待。

去年9月，中国国家主席习近平提出全球治理倡议，从奉行主权平等、遵守国际法治、践行多边主义、倡导以人为本、注重行动导向五个方面阐释重要内涵，为包括人工智能在内的新兴领域全球治理提供了根本遵循。

军事领域人工智能事关人类共同安危。中方一贯高度重视人工智能军事应用领域的风险防范和安全治理，主张：

一要坚持和平正义。不借助技术优势追求霸权、危害他国领土主权，维护战略平衡与稳定，推动人工智能军事应用服务于维护和平与安全。这是最根本的“负责任”。

二要坚持慎重负责。切实防范战略误判与冲突升级，要兼顾安全管控、技术发展及和平利用，分级分类管理，避免误用恶用，阻遏非国家行为体滥用。

三要坚持人道主义。确保有关武器系统始终处于人类控制之下，明确人是最终责任主体。使用中坚持区分、比例、

预防原则，严防无差别打击伤及无辜。

四要坚持多边协商。支持联合国发挥应有作用，推动构建各方普遍参与、具有广泛共识的治理框架。应避免泛化国家安全概念、人为制造科技壁垒。

中国代表团将就后续议题作进一步介绍。

Session 2 : Foundational terminology and concepts

当前人工智能向军事领域全面拓展，军事智能化加速演进，但发展趋势并不明朗，机遇与风险并存。军事人工智能是伴随军事智能化而生的新概念，横跨军事与人工智能两个领域，内涵丰富。界定“军事人工智能”或“军事智能”，既不能简单泛化为人工智能技术在军事领域的应用，导致扩大讨论范围，误伤民用领域人工智能技术；也不能狭义的限定为军事领域专用的人工智能技术，遗漏具有潜在安全风险的人工智能技术、产品与服务。

人工智能作为不断创新发展、概念层出不穷的新兴领域，学术界暂未就诸多数据形成定论或共识。以“致命性自主武器系统”（LAWS）为例，尽管该术语在 CCW 框架内已讨论近 10 年，但各方仍存分歧。可以预见，军事人工智能关键术语的讨论也将是一个漫长艰苦的过程。

为了提高多边讨论效率，推动各方尽早达成共识，中方建议：

一是在联合国机制框架内，初步拟定军事人工智能领域基础性和关键性的精简的术语清单，作为后续讨论的优先范围，避免议题过分发散。清单在后续讨论中可以进行动态调整，但一开始所列术语不宜太多。

二是按照自愿原则，各方（国家提案或联合提案）对清

单所列术语给出本国官方文件或权威文献或推荐认可的术语定义或解释。

三是优先对分歧较少或没有根本性分歧的术语，各方可达成非正式谅解，为日后此领域进程提供参考。

四是对于具有普遍共识，仅有少数国家存在分歧的术语，主要分歧方可提出书面文件予以说明，表达核心关切，表明可接受与不可接受的表述，提出修改意见，尽可能缩小分歧。

五是对于分歧较多的术语，可组织主要攸关方进行小范围磋商，通过对话对齐认识，消除误解，凝聚共识，力促形成可讨论的术语解释。

六是对于存在根本性的难以调和的分歧的术语解释，可采用各自表述的方式，均列入术语解释文件中。

Session 3: Current developments and trends in AI in the Military Domain

当前，军事人工智能已正式介入战争、进入战场，军事智能化不再是实验室中的算法、模型，正成为主导战争决策、作战计划、目标选择、作战效果分析等战争指挥的重要工具。无人作战集群、智能武器平台、智能辅助系统等，正在加速融入战场、嵌入作战链路、融入装备体系，也在影响军队组织编制、装备研发和保障模式。

面对军事人工智能带来的新挑战，中方认为：

一是要坚持以人为本的发展理念，军事智能化发展不能以全人类共同福祉为代价。坚持由人主导、人是最终责任主体这一根本性原则不能改变，确保有关武器始终处于人类控制之下。

二是确保人工智能军事应用安全可靠可控，应本着慎重负责的态度，在军事领域研发和使用人工智能技术，不能将战场作为试验场，把不成熟、不稳定的系统投入战争、积累数据。

三是要携手建立普遍参与的国际机制，坚持多边主义、开放包容原则，形成具有广泛共识的治理框架，不能以意识形态划线、搞“小圈子”和人为制造技术壁垒，要确保各国充分享有技术发展和应用权利。

Session 4: The Life Cycle of AI in the Military Domain (Part 1)

当前军事人工智能技术迭代加快，研发环节作为全生命周期的起点，是风险防控的第一道关口。中方认为，各国在开展相关技术研发时，应当平衡好安全与风险关切，坚守防御性导向。

在系统设计阶段，要合理设定技术边界。同时，国家应结合各自管理体制要求，根据智能系统技术应用特点，不断完善构建相应的评估与测试体系。

各国可围绕测试标准与安全规范等加强对话。各方可在自愿前提下开展技术交流与经验互鉴，避免将不成熟、未经验证的智能系统投入使用，为后续全流程治理筑牢基础。

Session 5: The Life Cycle of AI in the Military Domain (Part 2)

部署与使用是军事人工智能发挥作用的核心环节，应重视并保持人的有效控制。要根据不同作战场景、不同作战目标和作战环境，合理设置安全管控机制。

中方认为，人工智能军事应用不能也不应成为挑起地区矛盾、破坏国际和地区和平稳定、以大欺小、倚强凌弱的工

具。

近来，一些人工智能技术领先的国家出台政策，放松对军事人工智能的监管，加速部署军事人工智能系统，这将造成潜在的巨大风险。中方倡导各方以国际法为根本遵循，恪守区分、比例和预防原则，保护平民和民用设施，让军事人工智能成为维护稳定的工具，而非激化冲突的诱因。

Session 6:The Life Cycle of AI in the Military Domain (Part 3)

军事人工智能的销售、采购、跨境转让以及退役处置，直接关系全球防扩散体系安全。要强化尽职审查与风险评估，严格规范出口、采购及跨境转让行为，严防敏感算法、核心组件、关键数据流向冲突地区或非国家行为体。同时，中方坚决反对双重标准和技术滥用，以及搞排他性技术封锁。

当系统进入退役阶段，需确保数据安全，形成闭环管理，防范军事人工智能扩散风险。

Session 7:AI and Other Technologies

当前人工智能与核、生化、网络等敏感技术加速融合，催生复合型安全风险，给全球军控与安全治理带来全新挑战。中方对此高度关切。

各类大规模杀伤性武器、高危网络系统一旦与人工智能结合，会进一步放大误判风险、提升冲突破坏力。中方坚决反对利用人工智能赋能大规模杀伤性武器，严禁借助智能技术强化威慑、实施网络攻击。

国际社会应借鉴现有国际公约的成熟经验，将防扩散、保安全、促和平的原则延伸至技术融合领域。各方需加强风

险研判与协同管控，划定技术融合红线，完善配套监管规则。我们呼吁凝聚国际共识，构建复合风险防控体系，守住全球战略稳定的防线。

Session 8: International Cooperation & Capacity-Building in Artificial Intelligence in the Military Domain

军事人工智能技术不是少数人的蛋糕，更不是大国霸凌小国、弱国的工具和武器，全球治理也不应由部分国家说了算。相关新兴技术理应惠及所有国家，服务维护国际和地区和平与安全。

中方坚决反对在军事人工智能问题上恃强凌弱，把一个国家或几个国家制定的规则强加于人，搞针对特定国家的“小圈子”；坚决反对以意识形态划界建立歧视性壁垒，损害各国特别是发展中国家平等利用新兴科技的权利。“脱钩断链”“小院高墙”只会破坏国际经济秩序，扰乱全球产供链稳定。只有东西互鉴、共治普惠，寻求公约数、扩大合作面才是正解正道。

在人工智能领域发展不平衡，南北科技、数字和人工智能鸿沟呈扩大背景下，必须秉持国际合作精神，增强发展中国家代表性和发言权，落实好全球数字契约，让人工智能技术惠及广大发展中国家。在这方面，作为最具普遍性、代表性、权威性的政府间国际组织，联合国理应为军事人工智能全球治理挥应有作用，协调涉及人工智能发展和安全的重大问题。

中方愿积极支持联合国开展军事人工智能国际合作与能力建设相关工作，为推动建立普遍参与、广泛共识的治理规范贡献力量。

Session 9: Information exchange on State-led initiatives related to AI in the military domain

作为人工智能大国，中方一贯高度重视人工智能军事应用领域的风险防范和安全治理。在《全球人工智能治理倡议》以及相关立场文件中，中方从不同层面阐述了“以人为本的军事人工智能”理念。中方认为：

一是坚守慎重负责底线。各国尤其是大国在利用相关技术时，应慎重负责，摒弃绝对军事优势执念，防范战略误判与冲突升级，切实维护战略平衡与稳定，这是“负责任”核心要义。

二是秉持以人为本初心。恪守国际人道法，遵循全人类共同价值，尊重和保障人类尊严和人权。坚持由人主导，明确人是最终责任主体，确保有关武器系统始终处于人类控制之下。

三是恪守智能向善准则。各国应遵守相关国际法和公序良俗，推动人工智能军事应用服务于维护和平与安全，避免将之作为追求霸权、危害他国领土主权的工具。

四是落实敏捷治理原则。兼顾安全管控、技术发展及和平利用，加强前瞻性风险研判和人员培训，实施分级分类管理，避免误用恶用滥用，加强对黑客攻击、数据欺骗、恐怖组织滥用等各类风险管控。

五是坚持多边主义宗旨。支持联合国发挥应有作用，推动构建具有广泛共识的治理框架和标准规范。应避免泛化国家安全概念、人为制造科技壁垒，规则标准制定应由各国参与、共同商量。

中方将继续秉持开放、包容、互鉴原则，同各国加强沟

通交流，深入务实合作，推进以联合国为核心的人工智能多边治理进程，共同为军事人工智能全球治理做出贡献。

Session 11:Existing and Emerging Normative Proposals

关于军事人工智能治理机制，中方初步建议：

一是坚持自律原则。在新兴科技日新月异的背景下，军事人工智能发展具有不确定性，潜在风险具有不可预测性，失控危害具有较强的外溢性，安全治理具有滞后性。因此，军事人工智能治理首先是各国自身安全发展的刚性需要和优先责任。

中方呼吁国际社会加强军事领域人工智能安全治理能力建设，为军事人工智能全球治理探索可行的方案。包括但不限于发布政府官方文件，宣示本国军事人工智能发展理念、治理政策、核心立场、主要主张等；立法规范国内军事人工智能技术发展、装备研制及部署使用；建立技术标准体系，加强与完善人工智能“技术-产品-服务-基础设施”安全与应用的技术规范；建立测试评估体系，强化军事领域人工智能技术与产品的全生命周期安全、可靠、可信、可控测试评估；加强军事人工智能相关人员的培训，提高从业者的伦理意识、安全素养与治理能力。

二是坚持共商共建共享。以人工智能技术为核心的智能化技术与产业生态，是人类社会迈入智能化时代，赖以生存、安居乐业，健康发展的共同基石。人工智能技术缺陷，失效失控、误用滥用、发展失序是人类的共同敌人，发展安全、可靠、可信的人工智能公共产品，形成公平、合作、互信、良性竞争的 AI 产业生态是国际社会的共同期盼。我们坚持主

张构建以联合国为核心的全球人工智能治理机制，国家不分大小强弱，平等地参与人工智能发展与治理机制建设，帮扶欠发达国家技术研发、产业落地、安全治理能力建设，共享智能化的发展红利，破解发展难题，缓解南北差距，弥合代差鸿沟，助力构建和解、和平、合作、和谐的人类命运共同体。

三是兼顾公平与效率。当前智能化技术发展与产业应用日新月异，对人类社会生产生活、产业生态、战争形态影响日益深远，潜在风险如影随行，加快构建人工智能全球治理体系势在必行。我们注意到，当前一些多边机制陷入长期的争论而无果而终，既浪费了与会各方的大量外交资源，也错失了解决问题的最佳时机。建议可考虑在联合国机制下平台采用“小组攻研+大会决策”的运行机制，即针对军事人工智能治理相关的术语体系、治理机制、规约文件、技术标准等核心议题成立若干专业化、小规模、多方参与的工作小组，采用经常性的线上会议方式专题研究，形成成果文件草案；在此基础上，通过举行全体会议对各工作小组提交的成果文件进行讨论，形成共识的予以确认，存在分歧的交由工作小组继续研究完善，在不断凝聚与固化共识的基础上不断推进军事人工智能治理走深走实。

Session 12: Next Steps on AI in the Military Domain

在世界和平与发展面临多元挑战的背景下，各国应秉持共同、综合、合作、可持续的全球安全观，通过对话与合作，就如何规范人工智能军事应用寻求共识，构建有效的治理机制，避免人工智能军事应用给人类带来重大损害甚至灾难。

加强对人工智能军事应用的规范，预防和管控可能引发的风险，有利于增进国家间互信、维护全球战略稳定、防止军备竞赛、缓解人道主义关切，有助于打造包容性和建设性的安全伙伴关系，在人工智能领域践行构建人类命运共同体理念。为此，中方呼吁：

战略安全上，各国尤其是大国应本着慎重负责的态度在军事领域研发和使用人工智能技术，不谋求绝对军事优势，防止加剧战略误判、破坏战略互信、引发冲突升级、损害全球战略平衡与稳定。

军事政策上，各国在发展先进武器装备、提高正当国防能力的同时，应铭记人工智能的军事应用不应成为发动战争和追求霸权的工具，反对利用人工智能技术优势危害他国主权和领土安全的行为。

法律伦理上，各国研发、部署和使用相关武器系统应遵循人类共同价值观，坚持以人为本，秉持“智能向善”的原则，遵守国家或地区伦理道德准则。各国应确保新武器及其作战手段符合国际人道主义法和其他适用的国际法，努力减少附带伤亡、降低人员财产损失，避免相关武器系统的误用恶用，以及由此引发的滥杀滥伤。

技术安全上，各国应不断提升人工智能技术的安全性、可靠性和可控性，增强对人工智能技术的安全评估和管控能力，确保有关武器系统永远处于人类控制之下，保障人类可随时中止其运行。人工智能数据的安全必须得到保证，应限制人工智能数据的军事化使用。

研发操作上，各国应加强对人工智能研发活动的自我约束，在综合考虑作战环境和武器特点的基础上，在武器全生

命周期实施必要的人机交互。各国应始终坚持人类是最终责任主体，建立人工智能问责机制，对操作人员进行必要的培训。

风险管控上，各国应加强对人工智能军事应用的监管，特别是实施分级、分类管理，避免使用可能产生严重消极后果的不成熟技术。各国应加强对人工智能潜在风险的研判，包括采取必要措施，降低人工智能军事应用的扩散风险。

规则制定上，各国应坚持多边主义、开放包容的原则。为跟踪技术发展趋势，防范潜在安全风险，各国应开展政策对话，加强与国际组织、科技企业、技术社群、民间机构等各主体交流，增进理解与协作，致力于共同规范人工智能军事应用并建立普遍参与的国际机制，推动形成具有广泛共识的人工智能治理框架和标准规范。

国际合作上，发达国家应帮助发展中国家提升治理水平，考虑到人工智能技术的军民两用性质，在加强监管和治理的同时，避免采取以意识形态划线、泛化国家安全概念的做法，消除人为制造的科技壁垒，确保各国充分享有技术发展与和平利用的权利。