

全球人工智能安全治理需团结合作

贾平凡 唐雅晴

呼吁国际合作

《宣言》指出，人工智能带来巨大机遇，有潜力增强人类福祉、和平和繁荣，并已部署在日常生活的许多领域。但同时人工智能也带来重大潜在风险。《宣言》强调，人工智能的许多风险本质上是国际性的，因此最好通过国际合作来解决。与会国家和地区同意协力打造一个具有国际包容性的前沿人工智能安全科学研究网络，以尚未完全了解的人工智能风险和能力的加深理解。

《宣言》引发国际社会关注。据英国《金融时报》报道，《宣言》呼吁以“以人为本、值得信赖和负责任”的方式利用人工智能，是首个此类全球承诺。据《南华早报》报道，《宣言》提出双管齐下的议程：重点确定各国共同关心的AI技术风险，对其形成科学认知的同时，推动跨国政策形成以降低风险。

英国科学、创新和技术部大臣米歇尔·多内兰表示，此次峰会为围绕人工智能相关风险的全球讨论奠定了基础，也让国际社会意识到，必须加深对人工智能相关风险的理解并采取应对措施。

中国现代国际关系研究院科技与网络安全研究所研究员李艳表示，英国选择在布莱奇利园召开峰会意味深长，这里是二战期间盟军破译密码的主要地点。半个多世纪前，有“现代计算机科学与人工智能之父”之称的数学家图灵也是在此发明了第一代图灵计算机。今年以来，人工智能持续成为国际热点，人工智能发展与安全、国际竞争与规则制定牵动多方。在全球热议人工智能安全问题的背景下，英国借此时机召开峰会，其意有三：一是顺应国际大势，回应国际关切；二是搭建国际平台，彰显国家作为；三是成为未来国际合作重要贡献方和“斡旋者”。

“英国脱欧后，加大了对人工智能领域的投入，开发了许多底层技术和数据架构，人工智能产业发展迅速。”北京理工大学网络空间国际治理研究中心研究员裴轶说，这为英国举办此次人工智能安全峰会提供重要支撑。

根据英国政府发表的声明，下届人工智能安全峰会将于一年后在法国举办，韩国将在未来6个月内联合主办一次小型虚拟峰会。据西班牙《国家报》报道，此次峰会期间，英国建议政府之间成立一个人工智能小组，类似于处理气候变化问题的政府间气候变化专门委员会，每6个月举行一次会议，英国将分别与韩国、法国共同主办接下来两场。

中国科技部副部长吴朝晖率团参会并在峰会开幕式上发表讲话。中方指出，此次峰会为各国提供了人工智能安全与国际治理议题的重要对话平台与交流合作机会。人工智能治理攸关全人类命运，是世界各国面临的共同课题。发展人工智能，应当积极倡导以人为本，智能向善，加强技术风险管控，并在相互尊重、平等互利的原则基础上，鼓励各方协同共治，增强发展中国家在人工智能全球治理中的代表性和发言权，不断弥合智能鸿沟和治理能力差距。与会期间，中国代表团参与人工智能安全等问题讨论，积极推介中方提出的《全球人工智能治理倡议》，并与相关国家开展双边会谈。

环球热点

日前，为期两天的首届人工智能安全峰会在英国布莱奇利园举行。英国科学、创新和技术部的信息显示，本次峰会邀请了约100名各国官员、人工智能企业代表和专家。美国、中国、日本、德国、印度等20多个国家的政府代表以及联合国、经合组

织、国际电信联盟等多个国际组织的代表参会。各方代表就人工智能技术快速发展带来的风险和机遇展开讨论。28个国家和欧盟共同签署了《布莱奇利宣言》(下称《宣言》)，同意通过国际合作，建立人工智能(AI)监管方法。



在英国布莱奇利园，警察在首届人工智能安全峰会上执勤。新华社记者 李颖摄

共同应对风险

正如米歇尔·多内兰所言，世界正处于一场技术变革中，人工智能的蓬勃发展或将改变人类社会的方方面面，不仅带来巨大机遇，也会带来威胁全球稳定和破坏人类价值观的风险。抓住机遇、应对风险，全球要共同努力。

“人们很早就开始关注人工智能的风险问题。人工智能技术历史上历经几轮热潮，受益于大数据、算力等领域的突破，此轮人工智能发展浪潮最显著的特点是研发、应用和落地速度非常快，这就给平衡安全与发展的有效治理带来巨大挑战。”李艳说，最初，对人工智能风险的预判和管控主要聚焦于军事领域，如无人操作系统、无人指挥系统。随着ChatGPT等通用式人工智能技术突破，人工智能的应用场景愈加广泛，其风险也不再局限于某一特定领域。

在今天的联合国大会上，人工智能安全治理是一个新兴的热门议题。7月18日，联合国安理会举行主题为“人工智能给国际和平与安全带来的机遇与风险”高级别公开会。这是安理会在联合国总部举行的首次人工智能会议。联合国秘书长古特雷斯警告称，人工智能正在多领域发挥作用，但它也可能被怀有恶意的人使用，包括其用于恐怖主义、煽动仇恨和暴力等。他说，联合国的

人工智能监管机构应当规范、监督和执行人工智能规则，监管方式与其他机构监管航空、气候和核能的方式大致相同。

裴轶表示，目前人工智能技术主要造成三类问题。一是信息安全问题。人工智能需要实时更新海量数据用于模型训练。当前，在个人隐私领域中，相关企业、平台尚未达成遵循知情同意原则的共识，存在越权使用等现象。因此，用户如果缺乏相关意识，在使用人工智能技术时，很可能导致个人隐私数据泄露。二是虚假信息问题。尽管技术中立，但人为投放不实消息或片面事实可能导致人工智能模型产出虚假信息，也就是“深度伪造”。三是数字鸿沟问题。人工智能具备高效的信息检索和整合能力，虽然为人们的工作和生活提供了极大方便，但其产出的真实性和准确性并不牢靠。因此，信息甄别、自主决策能力强的使用者将更善于使用这项工具，而盲从的使用者则可能为其所害。随着两者使用技术的差距不断累积，数字鸿沟也就此加深。

李艳表示，目前，人工智能技术和产业尚处于发展初期，很多风险还难以评估。但如果人工智能安全风险明确显现或爆发后再制定治理规则，人类将付出巨大代价。国际社会需要给予人工智能一个密切审慎观察的空间，治理规则的制定需要在技术和社会的紧密互动中保持灵活性，随时根据实际情况校准政策，从而有效且前瞻性地解决问题。

兼顾安全与发展

为应对人工智能安全领域的风险，各国已纷纷展开行动。

10月30日，法国、德国、意大利三国在意大利首都罗马举行高级别会议并发表联合公报称，法德意三国正在制定人工智能战略愿景，商讨欧洲的联合项目和额外投资，并通过出台人工智能监管法规来管控风险。9月底，美国国家安全局宣布成立人工智能安全中心，负责监督美国国防和情报部门人工智能能力的开发和整合。6月，欧盟推出全球首部人工智能法案，对生成式人工智能企业提出更高的透明度要求。同月，英国首相苏纳克在访美期间和美国总统拜登发表《大西洋宣言》，同意在人工智能等领域加强合作。

此外，2019年，联合国教科文组织发布了《人工智能伦理问题建议书》，积极思考人工智能伦理问题，推动人工智能伦理原则在全球范围内的传播与实施。同年，二十国集团(G20)发布《G20人工智能原则》，强调人工智能的伦理和社会责任。2018年，欧盟委员会提出《人工智能的伦理准则》，强调人工智能的应用必须符合人类价值观和伦理原则，并建议采取措施加强人工智能的监管和透明度，促进其可持续发展。

作为联合国安理会常任理事国和国际社会负责任的一员，中国积极参与全球人工智能治理，致力于提升人工智能技术的安全性、可靠性、可控性、公平性。早在2021年12月，中国就向联合国《特定常规武器公约》第六次审议大会提交《中国关于加强人工智能伦理治理的立场文件》。这些主张坚持维护人类福祉，坚守公平正义，体现了大国责任担当，为解决全球人工智能治理难题贡献了中国智慧和方案。日前，中方在第三届“一带一路”国际合作高峰论坛期间提出《全球人工智能治理倡议》，就各方普遍关切的人工智能发展与治理问题给出建设性解决思路，为人工智能相关国际讨论和规则制定提供了蓝本。

“首届人工智能安全峰会为全球人工智能安全治理搭建了国际交流平台，引起了各方关切，为今后开展合作营造了良好的舆论氛围和国际政策环境。但此次峰会能否为全球人工智能治理搭建一个更具持续性、更有稳定性的机制，还存在较大不确定性，有待进一步观察。”李艳说。

“人工智能安全治理要达成全球性共识还面临不少挑战。”李艳表示，不同国家的国情和发展阶段不同，对人工智能安全治理的利益诉求不尽相同。当前，人工智能安全领域比较容易在原则和伦理层面达成共识，但是在军事、技术标准等具体领域达成共识的难度相对较大。但无论如何，国际社会就此保持积极的对话交流，总是有助于在发展和安全之间寻求最大的平衡。

裴轶表示，监管要兼顾安全与发展，通过细化立法领域、拓宽管辖范围、提高法律效力层级等方式，为人工智能的探索划定原则底线。

由迪拜商会主办的首届迪拜商业论坛日前成功举行。来自100多个国家和地区的2000多名各界代表围绕“变动中的经济力量：迪拜与全球贸易的未来”这一主题，就经济全球化、外商直接投资、数字化转型、新兴市场等议题进行了深入讨论。

今年初，迪拜提出“D33”经济议程，目标是到2033年实施100个转型项目，使迪拜的经济总量和对外贸易额均翻一番，并在未来10年内新增400个主要贸易伙伴城市。迪拜商业论坛作为实施“D33”经济议程的重要平台，旨在探讨合作、共谋发展。

论坛期间，迪拜商会同多个合作伙伴签署合作协议，其中包括阿联酋皇家贝尔特展览公司共同成立中国创新中心。这是中东地区首个专为中国企业展示创新技术与服务的平台。据介绍，迪拜商会中国创新中心将为中国企业展览展示、贸易代表团组织等提供支持，帮助中国企业特别是智慧城市、清洁能源、先进制造等领域的科技创新型企业进入迪拜市场。

“我们正在与来自200多个国家和地区的合作伙伴共同建设迪拜。”迪拜金融市场兼纳斯达克迪拜交易所首席执行官哈米德·阿里表示，近年来，越来越多中国企业来到迪拜投资兴业，迪拜可以助力中国企业更好开拓区域和国际市场，“我们也希望分享中国发展机遇，实现共同发展”。

迪拜商会总裁兼首席执行官穆罕默德·卢塔在接受本报记者采访时表示，通过加强经贸联系、促进可持续发展，各方能够共同应对挑战，提升民众生活水平。统计显示，中国作为迪拜最大贸易伙伴和最大投资来源国的地位不断巩固，已有超过8000家中资企业在迪拜投资兴业，仅今年上半年就有664家中国企业落户迪拜。卢塔说，迪拜将充分发挥自身优势，充当阿联酋与中国交流的重要桥梁，“阿中经贸合作潜力巨大。期待更多阿联酋企业到中国投资，为阿中可持续发展注入新动力”。

欧洲中央银行前行长让-克洛德·特里谢表示，地缘政治冲突、通胀高企、人工智能技术等对全球经济和金融转型带来深远影响。“面对挑战，各国只有加强合作、扩大开放，才能实现共同发展繁荣。”特里谢在接受本报记者采访时表示，中国采取一系列措施推动金融高质量发展值得赞赏。“中国金融业对外开放持续深化，中国稳慎扎实推进人民币国际化，有助于经济高质量发展。”

为阿中可持续发展注入新动力

本报记者 管克江 任皓宇

印尼努力改善城市公共交通

本报记者 曹师韵



11月5日早上，在雅加达哈利姆站候车大厅，人们刷票准备前往站台乘坐当日第一班高铁。新华社记者 徐钦摄

印度尼西亚首都雅加达首条轻轨正式通车以来，乘客数量稳步增加，积极作用不断显现。该轻轨线全长约41公里，连接雅加达市中心及周边卫星城市，全面运营后每天可运送乘客18万人次，有效提高首都和周边地区的通勤效率。

印尼人口总数约2.8亿，面临严重的城市交通拥堵问题。尽管印尼运行了快速公交、小型巴士、地铁等公共交通工具，但由于配套设施不完

善、线路较少、成本较高等原因，选择乘坐公共交通工具出行的民众并不多。印尼官方数据显示，私家车的持续增加导致雅加达及周边城市交通拥堵日益严重，并造成空气污染，每年给印尼造成的经济损失近100万亿印尼盾(1元人民币约合2137印尼盾)。

为破解交通拥堵难题，印尼近年来加大力度发展城市公共交通。以雅加达为例，该市目前拥有数十条较为完善的快速公交线路，使用红、蓝、

绿等不同颜色进行区分，采用快速、高频次的运营模式，并划定专用车道，与社会车辆隔离。2019年4月，雅加达首条地铁线路正式投入运营。2022年的一项调查显示，地铁开通后，最靠近地铁线路的地区空气污染水平下降27.4%。今年8月，印尼政府表示将为民众发放乘坐大雅加达地区轻轨、地铁及其他公共交通工具的补贴，以降低民众出行成本，鼓励更多人乘坐公共交通工具。

印尼同时致力于提高不同公共交通工具之间的接驳效率，例如加强火车站、轻轨站、快速公交站等站点之间的连通。雅加达还推出了非接触式智能卡和自动收费系统，应用于地铁、轻轨、专线巴士等。市民只需持一张支付卡，就可以乘坐所有交通工具。目前这一支付方式已覆盖雅加达80%的区域。

印尼还积极促进公共交通绿色低碳发展，计划在2030年实现全国公交电动化。按照规划，正在建设的印尼新首都80%的交通工具将实现电动化。印尼交通部副部长布迪·卡里亚·苏马迪表示，交通部欢迎公共和私营部门之间开展更紧密合作，共同创建更加安全、舒适、高效的公共交通系统。



环球掠影

近日，意大利“三色箭”飞行表演队喷出象征国旗三色的烟雾飞过罗马上空，庆祝国家团结日。新华社/法新