

首尔人工智能峰会宣言

首尔人工智能峰会 (AI Seoul Summit)

2024 年 5 月 21 日

中文翻译版 · 仅供参考，以英文原文为准

原文来源: https://www.mofa.go.kr/eng/brd/m_5674/view.do?page=1&seq=321007

参加首尔人工智能峰会领导人会议的与会者发表的《关于安全、创新和包容的人工智能的首尔宣言》

2024年5月21日

- 我们，代表澳大利亚、加拿大、欧盟、法国、德国、意大利、日本、大韩民国、新加坡共和国、英国和美利坚合众国的世界领导人，于2024年5月21日齐聚首尔人工智能峰会（AI Seoul Summit），确认我们致力于在人工智能（AI, Artificial Intelligence）史无前例的发展及其对经济和社会影响面前，促进国际合作与对话的共同承诺。
- 在2023年11月于英国布莱切利庄园（Bletchley Park）举行的人工智能安全峰会（AI Safety Summit）所奠定的工作基础上，我们认识到人工智能安全（AI safety）、创新（innovation）和包容性（inclusivity）是相互关联的目标，在关于人工智能治理的国际讨论中纳入这些优先事项非常重要，以应对人工智能的设计、开发、部署和使用所呈现和可能呈现的广泛机遇与挑战。
- 我们认识到人工智能治理框架之间互操作性（interoperability）的重要性，遵循基于风险的方法（risk-based approach），以最大化人工智能的利益并应对其广泛的风险，确保人工智能的设计、开发、部署和使用是安全、可靠和可信的。我们继续支持广岛进程国际行为准则（Hiroshima Process International Code of Conduct）对开发先进人工智能系统的组织

的运作化。我们认识到开发和部署前沿人工智能（frontier AI）的组织有特殊责任，并注意到前沿人工智能安全承诺（Frontier AI Safety Commitments）。

4. 我们支持本宣言参与方现有和正在进行的努力，以创建或扩大人工智能安全研究所（AI safety institutes）、研究项目和/或其他相关机构（包括监管机构），并努力通过培育这些组织之间的网络来促进安全研究合作和最佳实践分享。在此方面，我们欢迎附于本宣言的《关于人工智能安全科学国际合作的首尔意向声明》（Seoul Statement of Intent toward International Cooperation on AI Safety Science）。
5. 我们呼吁加强国际合作，推进人工智能安全、创新和包容性，以利用以人为本的人工智能来应对世界上最大的挑战，保护和促进民主价值观、法治和人权、基本自由和隐私，弥合国家之间和国家内部的人工智能和数字鸿沟，从而促进人类福祉的提升，并支持人工智能的实际应用，包括推进联合国可持续发展目标（UN Sustainable Development Goals）。
6. 我们倡导建立政策和治理框架，包括基于风险的方法，以促进安全、创新和包容的人工智能生态系统。框架应促进人类创造力与人工智能开发和使用之间的良性循环，促进社会文化、语言和性别多样性，并在商业和公共可用的人工智能系统的整个生命周期中促进环境可持续的技术和基础设施的开发和使用。
7. 我们确认积极的多利益相关方合作（multi-stakeholder collaboration）的重要性，包括政府、私营部门、学术界和公民社会，以培育安全、创新和包容的人工智能生态系统，以及跨境和跨学科合作的重要性。认识到所有国家都将受到人工智能利益和风险的影响，我们将在人工智能治理的讨论中积极纳入广泛的国际利益相关方。
8. 我们旨在通过参与联合国及其机构、G7、G20、经济合作与发展组织（OECD）、欧洲委员会和全球人工智能合作伙伴关系（GPAI, Global Partnership on AI）等其他国际倡议来加强人工智能治理方面的国际合作。在此背景下，我们认可广岛人工智能进程之友小组（Hiroshima AI Process Friends Group），欢迎最近更新的经合组织人工智能原则（OECD AI Principles），以及联合国大会最近通过的“抓住安全、可靠和可信的人工智能系统促进可持续发展的机遇”决议——该决议巩固了全球对人工智能系统需要保障措施以及开发、部署和使用人工智能向善的必要性的理解。我们欢迎在2024年9月未来峰会前就全球数字契约（Global Digital Compact）进行的讨论，并期待联合国秘书长人工智能高级别咨询机构（HLAB）的最终报告。
9. 认识到人工智能峰会对话作为推进人工智能治理讨论——促进人工智能安全、创新和包容——的高层论坛的价值，我们期待在即将在法国举行的第三次人工智能行动峰会（AI Action Summit）上再次相聚。

附件：《关于人工智能安全科学国际合作的首尔意向声明》