



人工智能：利益、风险 与规制

文/柯比·莱恩斯博士
伦敦国王学院战争研究系荣誉高级研究员

Dr Kobi Leins
Honorary Senior Fellow in the Department of War Studies,
King's College, London

从未知的能力到可信的工具，关于 AI 未来的发展，一切的根本在于问对问题。

日常生活中，关于人工智能（AI, artificial intelligence）最常见的对话是“我用 ChatGPT 做了某事”。而企业高管、政府和国际组织的议题却截然不同，他们探讨的是如何利用 AI 来减缓风险。

有人认为需要立即规制 AI，有人则将 AI 与反核运动相提并论，甚至还有人警告说 AI 将毁灭世界。同时，许多专家和新创企业则在努力让人们相信，AI 是帮助我们解决商业乃至个人问题的良计妙策，包括[爱、生命与死亡](#)。现在下结论还为时过早，更重要的是让合适的人进行充分的探讨。唯有如此，这项前所未有的技术才能够给人类发展提供支持和赋能。

问对问题

事实上，[许多关于 AI 的探讨](#)都并未触及关键的问题。其中包括更为广泛的社会性影响，如加剧社会的不平等，[将人简化为数据点](#)，以至于人可以被弃为冗余数据或不再具备价值。纵观历史，科学的每一次进步都伴随着利益和风险。事实上，我们可以以史为鉴，避免再犯同样的错误。尽管在某些方面有所不同，但伴随着传统的范式转变，AI 也带来了许多相同的潜在问题。过度承诺、淡化风险以及商业利益影响等问题屡见不鲜。目前有哪些新的进展？为什么我们要关注？

很多问题都是老生常谈。自从 20 世纪 50 年代魏森鲍姆（Weizenbaum）开创了神奇的语言模型研究并创造了第一个聊天机器人 Eliza，人们就一直关注着语言模型领域的进展。最近，有人提议可以将 ChatGPT2 用于自动司法判罚，甚至判处死刑，无需人工干预，这引起了数据科学领域的担忧。尽管目前这项技术得到了大规模数据集的加持，但许多老问题仍未解决。所以，新的进展体现在这些模型的速度和规模，以及数据的来源上。

实际上，我们不能依赖传统思维来看待 AI 的风险。

治理

好消息是我们已经有了整体的治理工具箱，其中包括与知识产权、企业行为、人权、歧视、合同与隐私等相关的国际和国家法律。来自全球的许多专家，比如爱德华·塞托（Edward Santow）教授，一直在呼吁提升律师技能，以便在其专业领域内更好地理解和应用现有法律和新技术。

与此同时，在法律基础上还需要考虑更多规制问题。目前已有不少规制框架，比如近期制定的欧盟 AI 法案、美国国家标准与技术研究院（NIST）和中国均出台了新的 AI 政策。但部分标准及法规仍存在漏洞，所以需要更新或修订。我们应该针对这些漏洞完善规制系统。

降低风险，实现利益最大化

实际上，我们不能依赖传统思维来看待 AI 的风险。美国亚利桑那州立大学（ASU）的安德鲁·梅纳德（Andrew Maynard）教授是经验丰富的风险专家，他坚定地认为：传统思维方式无法帮助我们“应对新的风险”。

国际标准，比如由 ISO/IEC 信息技术联合技术委员会人工智能分委会（[ISO/IEC JTC 1/SC 42](#)）制定的 AI 管理国际标准，有助于填补空白，健全规制。国际标准可以支持决策者和政策制定者制定一致的、可审计的数据和流程。这将在诸多方面为公司提升长期价值，包括环

境报告、利益相关方的可操作性与可信性。这也可以保证，在符合规制与其他治理工具要求的条件下，AI 带来的利益超过风险。

数据伦理的重要性不可忽视。如果应用得当，数据伦理可以帮助塑造团队意识——从领导决策到日常工作——不仅要“知其可为而为之”，更要“知其应为而为之”。

更重要的是，国际标准可以确保让合适的专家通过通用的语言来探讨关键性问题。创建符合需要的规制工具和文化可能要花费一些时间，但国际标准可以帮助我们更好地平衡其中的风险和利益。

关于柯比·莱恩斯

柯比·莱恩斯博士 (GAICD) 是人工智能、国际法和治理领域的全球专家。她关注数字伦理研究，举办过多次讲座，出版了多部著作，在推动人工智能的认知普及方面作出了突出贡献。她的著作将创新和实际应用贯通，把复杂概念拆解简化，方便多元化的受众群体理解。

柯比·莱恩斯将在 ISO 年会上发表讲话。欢迎您参与交流！加入我们的在线研讨“知或未知，AI 已至” (Ready or not, here comes AI)，了解人工智能带来的政治、社会和伦理影响。[点击此处注册](#)