

中国发展高层论坛 2024 年年会 “持续发展的中国”

人工智能发展与治理专题研讨会 (小组讨论一)

由国务院发展研究中心主办、中国发展研究基金会承办的中国发展高层论坛2024年年会，于2024年3月24-25日在北京钓鱼台国宾馆举行。3月24日下午举行了“人工智能发展与治理专题研讨会”，其中，“小组讨论一”环节的发言嘉宾有5位，分别为采埃孚集团董事会主席、首席执行官柯皓哲，Verimag实验室创始人、2007年图灵奖得主约瑟夫·希发基思，香港中文大学（深圳）前海国际事务研究院院长、广州粤港澳大湾区研究院理事长郑永年，麦肯锡公司全球总裁鲍勃·斯滕菲尔斯和标普全球集团总裁、首席执行官道格拉斯·彼得森。本场会议由国务院发展研究中心副主任张顺喜主持。

柯皓哲首先从汽车行业技术公司角度谈了其对人工智能快速发展的看法。他认为人工智能是一种转型性力量，正在改变着我们的工作方式，如工程师编码软件的方式、优化产品和流程的方式。他举四例加以说明。一是电动车发动机温度测量。电动车的发动机是高速旋转的，旋转快了以后就会有温度升高的问题，很多时候人工无法精确测量的转速可以改用人工智能进行测量，这对降低电机故障率有很大帮助。二是使用人工智能算法做自动驾驶。自动驾驶的车需要自己拐弯和掉头，需要知道在什么情况下做决策。现在训练的人工智能系统，可以验证车做出决定的准确性，从而满足安全行驶方面的要求。三是现阶段已经可以提供支持人工智能升级所需要的算力。四是人工智能

在公司管理中已经被广泛使用。

这四个例子充分说明，现在人工智能已经在公司运营中扮演着越来越重要的角色。柯皓哲就此表示，现在的挑战在于要培训公司员工使用人工智能，这实际上也是采埃孚努力的方向，让员工不要惧怕人工智能，能够使用人工智能让人类更加智慧地工作，而不是让人工智能替代人类。相信人工智能将成为汽车行业发展和运行不可或缺的一部分。政产学研用都要一起努力，找到一种安全的方式，以避免人工智能带来的风险。

约瑟夫·希发基思指出，当今世界，对人工智能的终极目标存在两种不同的观点。有些人认为它应该是通用人工智能，旨在构建超越人类思维能力的，在医疗保健、金融、交通运输和制造业等多个领域做出有效决策的超级智能代理。另一些人则认为，其终极目标是实现自主的人工智能，构建具有人类水平的智能机器，能够替代人类在现有组织中的工作，如自动驾驶汽车、智能工厂、自治的网络等等。但是目前我们所拥有的，还是一个弱人工智能，是智能助手，它通过与用户的互动来提供服务。约瑟夫·希发基思相信现在有三种不同的人工智能使用方式，包括作为助手、作为监视器、作为控制器来。中国可以利用自身庞大的工业系统，在迈向自主人工智能的道路上取得领先地位。

人工智能面临两种主要风险，一种是技术风险；另一种是人为风险，它可以通过监管或者法律的框架来控制。对由于人工智能所带来的自动化导致的失业问题，也可以通过适当的社会政策来应对。除此之外，还有两种社会性风险，一种是怎么在选择和性能之间取得平衡。如果我们不能确保系统是以公正、中立的方式使用可靠信息，则不要将决策权交给系统；另一种就是性能的提升。我们要考虑其是否与人类控制的缺失相平衡。

对于怎样建立一个全球性的监管框架，约瑟夫·希发基思表示，

人们都认识到必须要对人工智能进行监管，但对在实际中如何进行监管并没有达成一致。欧盟和美国之间，在人工智能监管方面也存在一定的差异。欧盟拥有更严格和更完善的人工智能监管法律，通过了《人工智能法案》和《数字服务法案》，采用风险管理的方法，要求人工智能系统在关键性的应用中具有高度的可靠性。而美国的监管相对较少有强制性，例如它发布的《人工智能行政命令》，就包括一些建议和指导性意见。美国的监管者和大型科技公司的立场是一致的，他们倡导符合伦理的人工智能原则。但是从技术角度看，它并不是基于非常严格的技术标准对人工智能系统进行评估。

约瑟夫·希发基思认为，目前联合国倡导的关于人工智能全球监管框架能够达成一致的机会并不大。美国正在利用其在生成式人工智能领域的优势，推行一种自我调节、以市场为中心的监管方法。但事实上，只有通过自主人工智能领域取得领先地位，并和相关国家联手建立安全的人工智能监管制度，才能够推动人工智能监管的大踏步前进。

郑永年指出，人工智能是个工具，发展与治理必须要同时发展、同时进步。人工智能领域的发展，主要有美国、欧盟、亚洲其他国家以及中国四个模式。美国是发展导向模式，缺少规制，生成式、商用的人工智能发展得很好；欧盟是规制导向的，主要把庞大的市场转化成为世界性的规则和规制；亚洲的新加坡、越南利用人工智能为经济发展服务，称为全国开放、有效管理的模式；中国主要是监管导向，人脸识别、可视化的图像处理、个人信息方面发展得不错。

在发展方面，中美两国在人工智能领域的发展各有比较优势，也是非常互补的，可以互相学习；在治理方面，更需要中美两国在人工智能治理领域进行合作。两国在人工智能领域都有不错的基础，合作的领域和可做的事情很多，合作将有利于世界的和平与发展。

鲍勃·斯腾菲尔斯围绕生成式人工智能的应用讲了四点看法。第

一，有巨大的潜力可以挖掘。麦肯锡通过研究人工智能科技的63个主要领域，未来规模能够达到2.6万亿-3.4万亿美元，制造、市场营销、软件工程、研发等四个领域将占到总规模的70%左右。第二，规模化是一个巨大的挑战。人工智能科技能够加速变革，需要我们提升自身技能。第三，对于企业而言，必须要重新改变组织结构的模式，包括改变评估人才的标准、缩短战略实施时间和人才储备时间，以及专注于客户和发展自己的技术人才等。第四，要平衡机遇和风险，尤其是全面评估风险，创建风险管理框架。在此基础上，为更好地应用人工智能，鲍勃·斯腾菲尔斯提出四个问题：一是你的企业是否已将人工智能融入发展路线图？二是是否已有相应的人才战略与发展战略相匹配？三是是否采取了过于激进的措施？四是是否把风险和机会放到同样的水平来思考？

道格拉斯·彼得森围绕标准普尔关于人工智能的落地执行，讲了五个主要元素。一是愿景。作为一家数据公司，人工智能将使标准普尔能够更加专注于决策和分析的层面，关注问题的核心，做核心的决策。二是数据。生成式人工智能取决于数据，标准普尔清洁数据、组织数据，从而可以让大模型去使用数据。三是人才。要保证我们的人也去学习，有机会了解人工智能，保证人类发展的方式是一样的。四是治理（组织结构）。标准普尔建立相应的AI架构，制定定期探讨制度，围绕人工智能来发展。五是安全性和质量。必须把控人工智能生产信息的质量，杜绝发布错误信息。

在此基础上，道格拉斯·彼得森还强调了人的重要性以及人工智能的治理。以标准普尔为代表的国际公司，既要有人工智能的全球性标准，还要通过合作，建立人工智能的治理架构，适应人工智能时代提出的新要求。

（中国发展出版社 撰稿：刘长杰、雒仁生、徐静；审稿：杨良敏
中国发展研究基金会 审校：丁孟宇）

—背景信息—

中国发展高层论坛是由国务院发展研究中心主办、中国发展研究基金会承办的国家级论坛。论坛秉承“与世界对话，谋共同发展”的宗旨，着眼于对外释放我国对外开放信号，自2000年创办以来，已发展成为服务中央决策和国家对外经济交往、促进中外企业交流合作的重要平台。

—媒体联络—

中国发展研究基金会

郭丝露 18666028168/64255855-8014
史雅帆 13810361966/64255855-8223
夏 天 18801375838/64255855-8086
施婉菁 18801090391/64255855-8090
邱凯娴 18301078627/64255855-8103



中国发展研究基金会
China Development Research
Foundation

中国北京东城区安定门外大街138号
皇城国际中心A座15层

邮 编: 100011
电 话: 86-10-64255855
传 真: 86-10-64255855-8100
网 址: www.cdrf.org.cn
电 邮: cdrf@cdrf.org.cn

Floor15, Tower a, Imperial International Center, No.138
Andingmen Wai Avenue, Dongcheng District, Beijing, 100011, China

Tel: 86-10-64255855
Fax: 86-10-64255855-8100
Website: www.cdrf.org.cn
E-mail: cdrf@cdrf.org.cn