



CISS 海外信息专报 (2024 年 7 月 11-12 日)

1、世界经济论坛：构建主权人工智能与本土人才以增强经济竞争力

7 月 4 日，世界经济论坛网站刊登 DAIMLAS 首席 AI 生态构建者约瑟夫·韦贝（Joseph Wehbe）的文章《构建主权 AI 与本土人才以增强经济竞争力》。文章认为，AI 技术的共性已形成，但完善生态系统需要集体努力，培养新一代 AI 人才来丰富本地生态。主权 AI 强调地方政府和联邦机构利用本土人才，根据本地政策和国家战略自主发展，超越对外部如硅谷的依赖，而这个过程需要硬件和基础设施支持。政府应重视本土人才，而非仅依赖资金，依据地区特色发展 AI，建立适应本地需求的生态系统。新兴技术中心应吸取硅谷经验并本地化 AI 发展。私营部门和中小企业提供项目和机会，是提升社区 AI 能力的关键。国家需迅速制定 AI 战略，避免在全球 AI 发展中落后，特别是适应生成式 AI 的发展趋势。其他国家可以学习沙特阿拉伯等 AI 战略成功的国家。政府利益相关者应认识到，AI 人才偏好有学习成长机会的生态系统。政府需创造并维护高质量的 AI 教育途径，建立本地 AI 生态系统，提升 AI 领导力和劳动力技能，引导 AI 应

用实践。政府和公共部门对 AI 的理解是培养 AI 人才和增强经济竞争力的基石。

<https://www.weforum.org/agenda/2024/07/sovereign-ai-talent-improve-economic-competitiveness/>

编译：刘嘉滨

2、大西洋理事会：各国人工智能安全研究所应加强合作

7月5日，美国大西洋理事会发表研究员考特尼·朗（Courtney Lang）的评论文章《推进人工智能安全需要国际合作》。文章认为，在首尔举办的人工智能安全峰会取得诸多成果，如各国同意在人工智能领域共享信息与资源、合作发展技术能力并进行人才培养、共同推进通用人工智能模型的测试评估。未来，各国人工智能安全研究机构还应采取以下步骤：首先，合作开发用于测试和评估人工智能模型的指标，并努力对风险进行量化评估。各国可以基于“红队测试”方法，开发新的“测试、评估、确认和验证（TEVV）”工具，旨在以国际一致的方法量化评估当前已存在的风险，并为潜在未来风险做好准备。其次，积极参与人工智能国际标准制定活动。支持国际标准化组织、国际电工委员会等机构当前有关人工智能标准制定的工作，争取以国际标准的形式确定人工智能测评流程。最后，与人工智能生态中的利益相关方广泛合作。各国人工智能安全研究机构不仅应加强与人工智能科研机构合作，还应将私营企业、公民社会、政府官

员等广泛主体纳入其决策框架。

<https://www.atlanticcouncil.org/blogs/new-atlanticist/advancing-ai-safety-requires-international-collaboration-heres-what-should-happen-next/>

编译：高隆绪

3、《福布斯》杂志：生成式人工智能的有效治理框架

7月9日，美国《福布斯》杂志刊登“数字盔甲”（Digital Armour）首席运营官玛尼·帕迪塞蒂（Mani Padisetti）的文章《释放人工智能的全部潜力》。文章认为，生成式人工智能可以提供令人难以置信的功能，但必须确保其合乎道德法律规范。因此，实施有效的人工智能治理框架至关重要。对此，文章提出五大建议：一是治理框架要确保明确的问责制，可以首先进行人工智能审计，以评估当前做法，然后建立明确的政策和程序，概述人工智能决策过程和结果。二是制定数据处理协议，确保敏感信息受保护，包括制定数据访问、存储和共享策略等。三是治理框架应包括定期监管系统是否存在偏见并采取纠正措施的机制，定义和维护道德标准，从而促进公平和公正。四是治理框架需要持续改进和风险管理。人工智能治理不是一次性的努力，而是一个持续过程。它涉及持续监控和评估人工智能系统，以识别和降低风险。这种积极主动的方法能够适应新挑战和机遇，确保人工智能工具保持有效和有益。

<https://www.forbes.com/sites/forbestechcouncil/2024/07/09/unlocking-the-full-potential-of-ai-why-oversight-and-responsibility-matter-with-third-party-tools/>

编译：曾星月

4、布鲁金斯学会：使国家人工智能研究资源推动包容性 AI

7月9日，美国布鲁金斯学会在其网站刊登其高级研究员马克·穆罗（Mark Muro）和布朗大学詹妮弗·王（Jennifer Wang）的文章《如何使国家人工智能研究资源中心推动包容性 AI》。文章认为，随着人工智能（AI）研究和开发（R&D）的影响日益显著，美国国家人工智能研究资源中心（NAIRR）的试点项目旨在解决 AI 研究资源获取不平等的问题。目前，先进计算资源（如图形处理单元 GPU）的高成本和稀缺性使得非精英学术机构难以进行大型语言模型（LLM）的查询、微调和训练，从而限制了创新速度和 AI 研究的多样性。NAIRR 试点项目通过提供对高级计算资源的访问，试图缓解这种“赢者通吃”动态，支持地理和人口多样性研究的包容性。此外，该项目还面临如何真正实现资源共享以及如何提高中低等研究机构 AI 研究和采纳技能的挑战，特别是在历史上被投资不足的黑人学院和少数民族服务机构中。本文认为，为了解决这些问题，NAIRR 试点不仅需要增加对计算资源的资金投入，还需要通过提供教育工具和资源、建立跨机构合作、并通过特定的研究竞赛和资助机会来扩大研究社区

参与。通过这些努力，NAIRR 试点可以促进技术和人力资源建设，支持学术界 AI 研究者，最终推动国家在 AI 研究领域地理和人口差异减少。

<https://www.brookings.edu/articles/how-the-national-artificial-intelligence-research-resource-can-pilot-inclusive-ai/>

编译：陆逸沛

5、CSIS: 校准北约对 AI 赋能决策系统的愿景

7月8日，美国战略与国际问题研究中心（CSIS）刊登其未来实验室博士后伊恩·雷诺兹（Ian Reynolds）和实验室副研究员亚西尔·阿塔兰博士（Yasir Atalan）的评论文章《校准北约对人工智能赋能决策系统的愿景》。文章表示，北约认为人工智能（AI）将在现代战争中确保决策的准确性和时效性。然而，北约在将 AI 技术整合到作战过程中仍面临若干挑战，具体包括盟友间 AI 可互操作性能力的差距、对 AI 解决战争迷雾问题的态度过于乐观，以及盟友在制定 AI 标准制定方面缺乏强制性机制。为了应对以上挑战，文章提出以下建议：首先，北约必须协调盟友间在 AI 赋能系统和数据共享机制方面的差异，北约数字政策委员会可以在该工作中发挥积极作用。同时，北约数据和 AI 审查委员会也应致力于协调盟友间的合作，建立数据集，以支持决策系统和辅助作战方案制定。其次，北约必须认识到，AI 并非解决政治和军事问题的“万金油”。开发人员和军事专家之间需保持

紧密合作关系，以确保人机交互的高效。此外，北约应通过公开对话达成共识，确保负责任使用 AI 的六项原则不损害成员国军事能力发展。最后，文章强调，北约在将 AI 应用于军事行动时，必须关注盟友之间在数据质量、模型训练以及标准制定方面的合作，客观审视 AI 技术对战争本质的影响。

<https://www.csis.org/analysis/calibrating-natos-vision-ai-enabled-decision-support>

编译：陈芳芳

6、布鲁金斯学会：AI 发展需要竞争与安全并重

7月8日，美国布鲁金斯学会官网刊登美国前联邦通信委员会主席汤姆·惠勒（Tom Wheeler）与非常驻高级研究员布莱尔·莱文（Blair Levin）的文章《对于 AI，既需要竞争，也需要安全》，探讨在 AI 发展过程中，竞争与安全并重的重要性。文章认为，尽管联邦贸易委员会（FTC）和司法部（DOJ）对 AI 公司的某些交易和合作进行反垄断调查是合理的，但 AI 的发展不应仅限于竞争，还应确保安全。竞争和安全并非互相排斥，监管机构应明确允许并鼓励在 AI 安全方面的合作。AI 技术的快速发展带来巨大的机遇和风险。制定并执行统一的安全标准至关重要，以避免企业为了利润而在安全方面偷工减料，导致社会承担不安全产品和服务的成本。AI 安全需要所有相关方的共同努力，特别是在开源 AI

模型广泛传播的背景下，这种合作显得尤为重要。文章还表示，仅靠自愿性的安全规范不足以应对广泛的安全问题，建议建立一个具有监督过程、市场激励和持续监督的 AI 安全模型。首先，建立一个由受监管的过程来识别问题并召集受影响公司和社会组织制定标准；其次，创建一个市场奖励机制，鼓励企业不仅达到基准标准，而且超越标准；最后，确保持续的合规监督，通过透明性来进行审计和强制执行。文章最后强调，AI 安全合作应该在不影响价格、产出或竞争强度的前提下进行，政府应为有利于公共利益的行业合作提供明确支持，而不会因为反垄断法的担忧而阻碍这种合作。尽管 AI 是新技术，但安全责任已经存在数百年，许多行业已经通过合作和政府监管建立了安全标准，如医学、工程和金融业。AI 安全的监管也应借鉴这些行业的成功模式，结合行业和政府的力量，制定和执行有效的安全标准。

<https://www.brookings.edu/articles/with-ai-we-need-both-competition-and-safety/>

编译：周正

7、兰德公司：人工智能公司并未真正落实安全优先政策

7月9日，美国兰德公司刊登兰德国土安全研究部管理、技术和能力项目（MTC）副主任、帕地兰德研究生院教授道格拉斯·杨（Douglas Yeung）的文章《人工智能公司并未真正落实安全优先政策》。文章认为，尽管科技公司致力于能

够在促进技术快速进步的同时保持其安全性，但由于群体思维局限性，科技公司在实践中往往难以真正落实安全政策。人工智能在带来便利的同时，也可能带来灾难性风险，基于此，OpenAI 等科技公司声称他们注重于保障 AI 产品安全性。但在实践中，群体思维让科技公司倾向于扼杀异议，传统的保密性文化依然大行其道，公司内部会无视安全风险警告并隐藏有关不安全做法的证据。对此，作者提出以下建议以真正落实科技公司对于人工智能安全性的承诺。首先，人工智能公司应禁止非贬低协议或保密条款，并鼓励提出不同意见；其次，学习诸如军事特种作战部队等高技能、以任务为中心的团队如何避免群体思维；最后，推动人工智能模型和公司内部运作更加透明。

<https://www.rand.org/pubs/commentary/2024/07/ai-companies-say-safety-is-a-priority-its-not.html>

编译：赵英慧

8、CSIS：伊朗新总统可能推动外交政策改变

7月10日，美国战略与国际问题研究中心（CSIS）发表评论员文章《伊朗新总统能否改变该政权的对抗性外交政策？》。文章认为，新当选的伊朗总统马苏德·佩泽什基安（Masoud Pezeshkian）可能为伊朗外交政策带来新变化。佩泽什基安在竞选中强调新的外交政策将优先寻求解除美西方制裁，这意味着与美国达成协议；然而，由于最高领袖哈

梅内伊的控制和强硬立场，变化可能是有限的。文章强调，伊朗外交政策的改变有着现实需求。首先，美西方对伊朗长期的制裁严重影响伊朗经济，寻求解除制裁已成为伊朗领导人优先事项。其次，此前的对抗性外交政策并未达到预期成效，前任总统莱希在执政后期也已开始推动改善与沙特和其他海湾国家关系，以缓解伊外交孤立。文章认为，美国大选结果也会对伊朗外交政策造成影响。若拜登连任或其他民主党领导人获胜，美国可能会推动与伊朗达成核协议，然而，实施这一政策将非常困难。如果特朗普再次当选，美国可能会减少对阿拉伯盟友的军事支持，但可能会推动改善美俄关系，从而减少俄罗斯与伊朗合作的可能。文章总结称，佩泽什基安的当选是伊朗向改变当前外交政策迈出的一步，然而，他的权力有限，同时不确定的美国政治也使得重大变化不太可能发生。

<https://www.csis.org/analysis/can-irans-new-president-change-regimes-confrontational-foreign-policy>

编译：孙至善

9、《金融时报》：印度为何会成为超级大国

7月9日，英国《金融时报》刊登专栏文章《印度为何会成为超级大国》。在2023年独立日的讲话中，印度总理纳伦德拉·莫迪勾勒了一个愿景，即印度有望在未来数十年内发展成与美国经济规模相当的超级大国。文章分析认为，

为实现这一目标，印度人均 GDP 年增长率需达到约 7.5%，尽管目标艰巨，但并非遥不可及。然而，印度面临严峻的外部环境，包括全球经济增长放缓、中美经贸摩擦导致的保护主义升温、人口结构变化以及长期的气候变化威胁，所有这些都可能阻碍印度经济腾飞。印度必须确保其经济增长速度至少是全球平均水平的两倍，这同样适用于其出口增长率，否则其经济将趋于封闭。同时，印度的战略位置和与西方国家的良好关系，尤其是在“中国+1”策略下，使其成为世界经济中的“连接国”。印度可以利用其庞大的侨民网络，尤其是美国的印度裔群体，推动经济多元化和产业升级。但是，印度面临的重大挑战在于内部，包括维持政治稳定、提升教育质量、巩固法治、改善基础设施、创造一流的投资环境、鼓励外国直接投资以及加速向绿色能源转型。最后，近期的选举结果预示着印度将继续保持政治稳定，这可能促使政府更加专注于经济和民生问题，避免卷入不必要的文化冲突。印度有潜力成为全球稳定力量，但实现超级大国地位的关键在于如何有效管理国内事务，实施促进增长的经济政策，并制定精明的外交策略，以应对复杂的国际形势和国内挑战。

<https://www.ft.com/content/7ebec65b-79fb-4366-8d55-bc817c8606ac>

编译：盛邵萱子

10、《政客》：中俄关系对西方意味着什么

7月10日，美国《政客》网站发布瑞典国际事务研究所斯德哥尔摩东欧研究中心（SCEEUS）分析师雨果·冯·埃森（Hugo von Essen）和安德烈亚斯·乌姆兰（Andreas Umland）的文章《中俄关系对西方意味着什么》。作者认为，近年来中俄关系日益密切，这使俄罗斯在经济方面得以继续维持。但是，中俄在几方面的利益仍存在分歧，如西方延长对俄罗斯银行和企业的制裁，使中国对俄罗斯的投资面临风险；2024年3月和4月，中国对俄出口首次同比下降，反映出俄罗斯进口商因美国施压而面临的付款问题等。尽管当前中俄关系趋于紧密，但这个非正式“联盟”出现裂痕的可能性也很明显。如俄罗斯与朝鲜日益密切的关系可能会导致三国之间出现关系危机。在这种背景下，西方保持决心和团结是很重要的，这意味着向乌克兰提供它所需要的一切支持。在西方支持下，乌克兰在战场上取得胜利将是最佳应对方式。

<https://www.politico.eu/article/china-russia-alliance-global-trade-economy-war-ukraine-beijing-moscow-liberal-front-taiwan-banks/>

编译：杨奕萌

11、传统基金会：欧洲盟友的安全政策与美国预期不符

7月8日，美国传统基金会刊登其国防预算政策顾问、艾利森国家安全中心政策顾问威尔逊·比弗（Wilson Beaver）撰写的文章《欧洲盟友的安全自主性存在分化》。文章认为，

当前美国军费开支占 GDP 比重远低于整个冷战时期，而美军却需要在多个地区维护所谓“国家利益”。面对多极化世界和国内严重的通胀问题，美国需重新分配其战略资源，使战略重心从欧洲调整至“印太”地区。文章的主要观点为：第一，美国的优势在于以美国为中心的盟友体系，但是其盟友却长期依赖美国提供安全保障，尤其是欧盟国家更愿意将钱花在社会运动与气候治理等非安全议题上。第二，欧洲各国对乌军事支援存在分化，比如德国、英国、波兰等行动积极，法国的实际援助则与承诺不符，而且远远低于其能力，使美国不得不填补缺位。第三，文章呼吁，美国必须要求欧洲盟友为自身安全承担更多责任并且加大对乌援助，使美国有余力在“印太”地区围堵中国。

<https://www.heritage.org/defense/report/some-european-allies-have-taken-responsibility-their-own-security-others-have-not>

编译：王昕怡

12、《报业辛迪加》：新左派需要什么

7月10日，《报业辛迪加》网站刊登哈佛大学肯尼迪学院国际政治经济学教授丹尼·罗德里克（Dani Rodrik）撰写的文章《新左派需要什么》。文章分析当前世界左派政党处境。首先，近年来法国和英国的选举结果以及美国的总统竞选，都反映了左翼政党在重新定义自身身份并提供极右翼替代方案时所面临的挑战。2008年全球金融危机后，极右翼率

先利用对新自由主义和超全球化的反感情绪。十年前“左派的退位”是普遍共识，当前左翼政党状况已有所改善，但仍面临许多问题。再次，左翼若想重新获得政治力量必须回归根基，重新代表劳动人民利益。右翼民粹主义崛起的核心在于经济不安全感增加、中产阶级受到侵蚀及落后地区好工作消失。左翼需应对这些趋势，提出可信的替代方案。文章总结称，新左派必须直面新的经济结构和提高生产力的必要性，只有这样才能成为未来真正的政治力量，成为极右翼的可靠替代品。

<https://www.project-syndicate.org/commentary/economic-program-that-can-beat-the-populist-right-by-dani-rodrik-2024-07>

编译：叶如静

13、布鲁金斯学会：围绕竞争和安全构建人工智能的未来

7月8日，美国布鲁金斯学会网站刊登学会治理研究访问学者汤姆·惠勒（Tom Wheeler）和大都会政策项目非常驻高级研究员布莱尔·莱文（Blair Levin）共同撰写的文章《围绕竞争和安全构建人工智能的未来》。文章提出，对人工智能的监管必须鼓励在人工智能安全方面的合作，同时又不能促成反竞争联盟。围绕竞争和安全构建人工智能的未来，挑战在于确保看似人工智能安全技术合作的行为不会掩盖可能存在的反竞争联盟。当前自愿制定的人工智能安全准则存在两个根本问题。一方面，完全自愿的安全方法没有足够广

度解决人工智能生态系统中的广泛问题，但现实要求人工智能生态系统中的所有环节都达到保护所有人的最低标准。另一方面，从定义上讲，自愿意味着缺乏执行力。为此，监管部门必须通过创建人工智能安全模型弥补自愿承诺之间的差距，人工智能安全模型包含三个基本组成部分。首先，制定标准的监督过程，目的是识别问题并召集受影响的公司和民间社会制定标准。其次，建立奖励超标企业的市场，引导市场激励达到基本标准甚至超过标准的企业，形成良性循环。最后，多个利益相关者（包括政府）合作持续监督合规活动。此外，这种新的人工智能安全模式也应该借鉴诸多行业与政府联盟的范例，以制定可执行的标准。

<https://www.brookings.edu/articles/with-ai-we-need-both-competition-and-safety/>

编译：宋琳琳

编译：孙至善、盛邵萱子、杨奕萌、王昕怡、叶如静、宋琳琳、刘嘉滨、赵英慧、周正、曾星月、陈芳芳、陆逸沛、高隆绪

审核：贺刚、申青青、张丁、郑乐锋