最终报告

美国人工智能国家安全委员会

美国人工
智能国家
安全委员会

委员会成员

主席：埃里克·施密特　　副主席：罗伯特·沃克

萨弗拉·卡茨　　埃里克·霍尔维茨

史蒂夫·钱　　安德鲁·贾西

米尼翁·克里伯恩　　吉尔曼·路易

克里斯·达比　　威廉·马克

肯尼斯·福特　　杰森·马瑟尼

若斯-玛丽 葛丽芙　　卡塔琳娜·麦克法兰

安德鲁·摩尔

主席和副主席寄语

目前，美国民众尚未充分认识到人工智能革命对国民经济、国家安全和人民福祉的深远影响。在人工智能技术的威力和局限性方面，我们还有很多需要学习和了解的地方。尽管如此，我们仍然需要进行重大决策：如何通过加快人工智能创新，造福美国并对人工智能的恶意用途进行抵制和防御。

在进行决策时，我们的领导人遭遇了亨利·基辛格先生描述的经典治国困境："当你的行动范围处于最大水平时，你采取此项行动所依据的知识总是处于最低水平；而当你的知识范围处于最大水平时，你的行动范围往往消失的无影无踪。"当然，就目前而言，美国的行动范围仍然存在，但是回旋余地却正在缩小。

人工智能国家安全委员会（NSCAI)是一个由 15 名专家组成的国会两党委员会，成员包括技术人员、国家安全研究人员、企业高管和学界领袖。自从组建以来，NSCAI 一直在向外界发出警告：在人工智能时代，美国在防御层面或竞争层面，尚未做好准备。这是我们必须直面的严峻现实，它要求美国必须举全国之力，采取全面行动。在最终报告中，我们提出了防御人工智能威胁和以负责任的态度在国家安全领域应用人工智能的策略，并且指出了在这场关乎经济繁荣、国家安全和人民福祉的全面技术竞争中的获胜之道。美国政府无法单独完成上述任务，它需要与来自产业界、学术界和民间的坚定合作伙伴展开全方位合作。与此同时，为了在人工智能时代打造一个更加安全自由的世界，美国需要争取铁杆盟友的支持，并积极开发全新的合作伙伴。

人工智能是一项激动人心的技术，它是几代以来普惠人类的最强大工具。目前，科学家们利用人工智能技术，已经在生物、医药和天体物理等领域取得了惊人的进展。这些进展不是科技博览会上博取观众眼球的实验，它们真正提高了人类的生活水平，并且为我们逐渐揭开了自然界的神秘面纱，因此人工智能是当之无愧的"改变游戏规则"的伟大发现。

人工智能系统也被用于构建军事力量。我们担心在未来的武装冲突中，人工智能工具将成为人类首选的武器。而且，由于人工智能具备双重用途，通常以开源的形式存在并且传播速度极快，因此它不再是超级大国的专利或科幻小说中的奇思妙想。反政府势力已经开始利用人工智能进行虚假信息活动，试图在民主国家制造分裂和阻止我们对现

实世界的认知。政府、罪犯和恐怖分子将利用人工智能进行网络攻击，并且通过人工智能软件和商业化无人机的结合来打造“智能武器”。目前，美国的军事对手正在积极进行人工智能概念和平台的整合，以此挑战美国长达数十年的技术优势，而这一点已经不再是秘密。在缺少基于人工智能的全面军事能力和新型作战方式的情况下，我们无法抵御人工智能带来的威胁。我们希望国家安全部门和联邦政府机构能够接触并使用这个星球上最先进的技术，从而有效保护自身安全和美国民众的安全，以及保护自身的利益及盟友与合作伙伴的利益。

尽管我们已经开展了一系列激动人心的实验并启动了若干小型人工智能项目，但是美国政府距离真正“为人工智能时代的到来做好充分准备”还有相当大的距离。人工智能国家安全委员会中的商业领袖们对于美国政府在人工智能领域的缓慢进展倍感失望，因为他们深知大型机构可以部署和使用人工智能。诚然，在任何一个行业推动人工智能整合都是一件困难的任务；特别是在国家安全领域，我们常常面临着一些特殊挑战。但是，那些心怀坚定信念的领导人可以推动变革。我们需要五角大楼和联邦政府领导人积极打造技术性基础设施，将理念和实验与全新概念和操作进行有机结合。到 2025 年，国防部和情报部门必须实现“人工智能就绪”状态。

我们应当拥抱人工智能竞争。目前，这种竞争已经催生了对数据、计算能力和关键性元素——能够进行人工智能攻关的稀缺人才的巨大需求。人工智能可以横跨多个技术领域并在众多行业发挥作用，这充分证明了人工智能的威力不可小觑，同时也说明了另一个关键性事实：人工智能已经成为更广泛的全球性技术竞争的一部分。竞争将加速创新。当人工智能用于登月计划等造福人类的伟大项目时，我们应当像昔日研发疫苗那样，与合作伙伴携手前行。而且，更重要的是，由于人工智能竞赛正在加剧中美两国之间的战略竞争，因此我们必须赢得人工智能竞赛的胜利。美国民众应对中国的人工智能规划、资源和进展保持密切关注。与此同时，我们必须承认：在人工智能的部分领域内，中国是一个旗鼓相当的对手；而在人工智能的若干具体应用上，中国甚至是世界公认的领袖。在下一个十年内，中国将超越美国成为全球人工智能的霸主，对于这个东方大国的勃勃雄心，我们必须严肃对待。

人工智能竞赛也是一场价值观的竞赛。对于世界上任何一个珍视个人自由的民众而言，专制独裁政府对于人工智能的使用让人不寒而栗。它将人工智能作为压迫和监视的工具，这种现象不仅在其国内司空见惯，在全球范围内也愈演愈烈，而这与我们一直以来坚信的人工智能的用途背道而驰。人工智能的未来可能会是民主的，但是我们已经充分了解到技术可以使一个国家对外加强威权主义，对内实施极端主义。因此，我们不能想当然地认为未来的技术趋势将强化民主制度，而不会动摇它的根基。我们必须与其他民主国家和私有部门进行合作，共同制定人工智能技术的隐私保护标准，推动民主规范，

指导人工智能的使用，从而确保民主国家能够负责任地将人工智能工具用于维护国家安全。

鉴于竞争的巨大利害关系，我们重点建议美国在下列方面采取行动：

领导层

归根结底，我们有义务说服美国政府领导人：为了赢得人工智能时代，他们必须做出艰难决定，并支付人工智能研发的先期投资。在美国，责任止于总统，人工智能战略始于白宫。二战结束之后，我们组建了国家安全委员会来处理战后面临的各种挑战。今天，我们需要组建技术竞争力委员会，并授权委员会制定和出台各项战略，应对人工智能和相关技术面临的复杂安全、经济和科技挑战。这种领导责任应推广至所有关键性国家安全部门和机构。

人才

人才缺口已经成为美国政府最明显的人工智能短板，也是美国政府采购、研发和部署用于国家安全的人工智能时所面临的最大一项障碍。有些人认为，只需在国家安全部门或机构内为硅谷的技术专家们新增加若干岗位，一切就将万事大吉。但是，事情并非如此简单。我们需要组建全新的数字服务学院和民间人才国家储备，将技术人才培养上升到与军事人才培养相同的战略高度。总的来说，数字时代，美国需要组建一支专业化的数字队伍。与此同时，美国需要改善 STEM 教育（科学、技术、工程和数学）以及高技术移民引进和留用制度，从而赢得国际人才竞争。

硬件

微电子技术为人工智能的发展提供不竭的动力。目前，美国已经不再是全球最精密芯片的生产商。我们不想夸大美国科技地位的不稳定性。但是，鉴于绝大多数先进芯片都是在与我们的主要竞争对手相距仅仅 110 英里（175 公里）水域之外的一家工厂内进行生产，我们必须重新评估供应链的弹性和安全性。最近，汽车制造业芯片短缺造成一家美国汽车公司损失预计高达 25 亿美元。相比较而言，战略封锁造成的损失将更加巨大，并且严重威胁国家安全。因此，启动产业振兴项目，向国内芯片制造业提供大约 350 亿美元的联邦投资和激励计划，无疑是一笔划算的买卖，否则美国将不得不依赖其他国家制造的发动机为能够塑造未来的机器提供动力。

创新投资

我们担心只有少数几家大型企业和实力强大的州具备实现人工智能巨大突破所需的资源。尽管开源工具已经得到普及，但是在创新前沿，对计算能力和海量数据的需求却不断飙升，因为它们是改善算法不可或缺的元素。联邦政府必须与美国企业进行合作，确保美国的领导地位，支持多样化人工智能应用的开发，在最广泛意义上促进美国的国家利益。与初次报告和中期报告不同，我们在最终报告中下调了美国政府需要进行的投资金额。我们建议联邦政府先期拨款 400 亿美元，用于扩大联邦人工智能研发规模，在民间普及人工智能研发活动，以此提高在未来取得人工智能突破的几率。我们还需要在美国境内建造安全的数字基础设施、共享式云计算接入系统和“智慧城市”，使人工智能能够真正造福美国人民。我们预计在未来的几年内，联邦政府还需追加数千亿美元的投资。

现在，我们不应对产业政策进行抽象批评或担心支出赤字会阻碍前进的步伐。1956 年，德怀特·艾森豪威尔总统，一位财政保守的共和党人，与民主党占据多数席位的国会开展合作，拨款 100 亿美元在全美境内建造州际高速公路。按照今天的物价指数进行计算，这笔拨款相当于 960 亿美元。当然，我们可以效仿他们的做法，对这个国家的未来进行投资。

我们对于委员会中的两党合作倍感自豪。我们共同探讨、相互学习，并在关键问题上达成共识。我们很荣幸能够向国会和总统献计献策。用温斯顿·丘吉尔的话说，竞争的大幕已经拉开，这场竞争关系到美国的经济繁荣、国家安全和人民福祉。我们在报告中阐述了美国在人工智能时代开展防御、进行竞争和获取胜利应采取的首要措施。

主席：埃里克·施密特

副主席：罗伯特·沃克

执行总干事寄语

竞争伊始，美国更新

两年前，在人工智能国家安全委员会组建之初，我们对于这项工作的内容和前景几乎毫无头绪。不过，很多朋友和盟友倒是对我们寄予厚望，他们希望我们能够出色完成国会交付的任务——保持美国在人工智能领域的优势。

一路走来，我们得到了美国政府各大部门和机构的支持。许多部门和机构还向我们提供了大量资源，包括详细说明文职人员和军事人员的情况，并花费大量时间帮助我们了解它们的任务和优先事项。国会议员和工作人员与我们密切合作，共同推动政府在国家安全领域采用人工智能。

在执行委员会的工作过程中，我们与来自私人部门、学术界、民间团体和政府机构的数百名代表进行了接触，收到了大量机密和非机密性简报。与此同时，部分思考、应用和开发人工智能的各界人士在百忙之中抽出宝贵时间与我们进行了广泛的交流。

我们与几乎所有的合作伙伴在以下三点达成共识：第一，人工智能是一项非常强大的技术；第二，加大人工智能创新投资已经时不我待；第三，应在坚持自由原则的基础之上，开发和使用人工智能。

我们还与新老盟友进行了会晤。从印度新德里、以色列特拉维夫到英国伦敦，盟友们表达了与美国共同加深人工智能领域合作的强烈意愿。

在此，我谨代表人工智能国家安全委员会，对于向委员会提供服务的志愿者、在委员会内工作的实习生、与委员会分享经验和洞见的专家学者，以及与委员会建立深厚友谊的各界人士表示诚挚的谢意。我还要特别感谢委员会内的一群富有牺牲精神的全日制工作人员，他们当中的很多人辞去了重要职务加入我们，与我们共同执行这项重要的基础性工作。

在过去的两年里，我们承载了社会各界的期盼：人工智能可以给国民经济、人民福祉和国家安全带来巨大利益。与此同时，社会各界也表达了关切和忧虑：与其他技术一样，人工智能会带来全新挑战并加剧现有问题。我们认真听取、严肃对待这些意见，包括各种质疑。

最终，我们出色地完成了国会交付的使命，并且得出宝贵的结论：如果美国能够在坚持价值观的基础上接受并投资人工智能，它将改变美国的未来，确保美国及其盟友能够继续改造世界，普惠人类。

谢谢！

伊利·巴杰拉克塔里

执行摘要

在谈到人工智能对国家安全的影响时，我们无法找到与其对应的历史参照物。人工智能并不像蝙蝠翼隐形轰炸机那样是一项单独的技术突破。人工智能霸主地位的争夺也不像太空登月竞赛。甚至，它与电力等通用技术也无法进行比较。但是，托马斯·爱迪生对电力的描述，也同样适用于人工智能的未来："它是所有领域的核心……它拥有改变人类生活的秘密"。爱迪生的惊人评价来自于他的谦逊——用他的话说，"与出现的可能性相比，我的发明微不足道。"

人工智能国家安全委员会认为，关于人工智能及其未来的应用，还有很多可能性有待发现。不过，凭借以下两个已知的事实，我们仍然可以对人工智能的发展有所了解，并做出预判。

首先，计算机系统解决问题和执行任务的能力正在飞速提高，在某些情况下甚至超过了人类的表现。人工智能技术是近几代以来最强大的工具，它可以丰富知识、促进繁荣、增加人类经验。人工智能技术也是典型的"两用"技术。机器拥有比人类更快速准确的感知、评估和行动能力，这在任何领域无疑都是极大的竞争优势——无论是民用还是军用。对于掌握人工智能技术的公司和国家来说，它将是巨大的力量源泉。

其次，人工智能正在扩大美国已经显现的颓势。这是自二战以来，美国的技术优势首次受到威胁，而这恰恰是其经济和军事实力的支柱。如果目前的趋势不发生改变，中国将凭借其强大的实力、人才和雄心在未来十年内超越美国，成为人工智能领域的世界霸主。与此同时，人工智能正在加深网络攻击和虚假信息活动带来的威胁，俄罗斯和其他国家正在利用这些攻击和虚假信息活动渗透我们的社会，窃取我们的数据，干预我们的民主，而这不过是冰山一角。同时，新冠肺炎肆虐和气候变化所体现的全球危机，凸显了扩大国家安全概念和寻找创新技术解决方案的必要性。

“人工智能国家安全委员会最终报告提出了一个综合性的国家战略，旨在重整政府、重新定位国家并团结我们最亲密的盟友和合作伙伴，在即将到来的人工智能加速竞争和加剧冲突的时代进行防御和竞争。”

鉴于以上事实，委员会得出结论：美国必须立即行动起来，积极部署人工智能系统，加大对人工智能创新的投资，从而保护国家安全、促进经济繁荣、保障民主的未来。今天，在与信念坚定的对手进行的这场技术竞争中，美国政府尚未为夺取最终的胜利进行任何组织动员和投资，也未对有效抵御人工智能带来的威胁和快速在国家安全领域应用人工智能进行任何准备。今天，不能只是简单地逐步调整联邦研究预算，或者简单地为硅谷的技术专家们在五角大楼增加一些新职位。美国政府需要在心态上作出重大改变。今天，我们需要白宫的领导、内阁成员的行动以及国会两党的支持，才能赢得人工智能时代的胜利。

人工智能国家安全委员会最终报告提出了一个综合性的国家战略，旨在重整政府、重新定位国家并团结我们最亲密的盟友和合作伙伴，在即将到来的人工智能加速竞争和加剧冲突的时代进行防御和竞争。这将是一套双管齐下的方法。第一部分“在人工智能时代保卫美国”概述了利害关系，阐述了美国在抵御与人工智能相关的一系列威胁时必须采取的措施，并就如何负责任地使用人工智能技术保护美国人民和利益给出建议。第二部分“赢得技术竞争”讨论了人工智能竞争的关键要素，并对政府必须采取的行动给出建议，从而促进人工智能创新，提高国家竞争力，保护美国的关键优势。值得注意的是，这些建议中的各项活动是有机结合的整体，因此政府必须全面采取这些行动才能最大程度发挥其效力。

第一部分：在人工智能时代保卫美国

美国的战略竞争对手已经开始开发用于军事和其他恶意用途的人工智能技术，从换脸工具“深度伪造”到廉价的民用杀人无人机，人工智能给流氓国家、恐怖分子和罪犯提供了新工具。在这种情况下，人工智能技术强化的新型军事能力将成为全新武装冲突时代人类首选的作战工具。为了保障国家安全，美国必须迅速以负责任的态度采用人工智能来抵御这些威胁。否则，赤手空拳对阵以机器速度运行的具有人工智能能力的对手，将打开灾难的大门。如果没有人工智能的协助，人类将无法有效对抗或抵御人工智能支持的网络攻击、虚假信息攻击、无人机集群攻击或导弹攻击。国家安全事务人员必须接触并使用这个星球上最出色的技术来保护自己、践行使命和保卫人民。委员会建议政府采取下列行动：

抵御人工智能对美国自由开放社会的新威胁。各行各业的数字依赖正在将个人和商业脆弱性转化为潜在的国家安全弱点。今天，敌对分子利用人工智能加强虚假信息战和网络攻击。通过收集美国民众的数据，敌对分子正在建立关于美国民众信仰、行为和生物数据的数据库，以便有针对性地操纵或胁迫个人。这种数据收集风暴带来的国外势力影响和干涉需要我们进行一系列组织和政策改革，增强我们的工作弹性。政府需要成立一个特别工作组和 7x24 小时行动中心来对抗数字虚假信息，需要更好地保护自己的数据库，并在审查外国投资、管理供应链风险和保护数据立法中优先考虑数据安全。政府应该利用人工智能网络防御来防止人工智能网络攻击。生物安全必须成为国家安全政策的首要优先事项。

为未来的战争做好准备。如果我们的军队不能加速采用人工智能，他们可能会在未来十年内失去军事技术优势。这需要将自上而下的领导能力与自下而上的创新结合起来，从而确保与军事行动相关的人工智能应用能够及时到位。为此，国防部应该：

首先，到 2025 年为人工智能的广泛整合奠定基础。这包括建立共同的数字基础设施，培养有数字知识的劳动力人才，建立更灵活的采购、预算和监督流程。此外，它还需要从战略上剥离缺乏人工智能操作的作战系统，并在下一代军事能力上进行投资。

其次，到 2025 年实现军事人工智能战备状态。五角大楼领导人现在必须采取行动，推动组织改革，创新作战概念，建立人工智能和数字战备状态绩效目标，确定联合作战网络架构。国防部必须扩大人工智能研发活动组合的范畴，对特定领域给予持续性聚焦和关注。它还需要提升与盟友和合作伙伴的人工智能互操作性。

管理人工智能和自主武器相关的风险。人工智能将使武器系统具有新的自主性能。与此同时，这也引发了一系列与使用致命武力相关的法律、道德和战略问题。当人类指挥官或作战人员授权使用人工智能时，我们应以符合国际人道主义法的方式使用经过准确设计和严格测试的人工智能和自主武器系统。国防部应严格审查现有武器系统和定位程序，包括自主武器系统专用协议以及强有力的人工智能道德约束原则，确保美国部署安全可靠的人工智能和自主武器系统并以合法方式进行使用。寻求在全球范围内禁止使用人工智能和自主武器系统既不可行，也不符合美国利益。但是，在未经审查的情况下，在全球范围内使用此类系统将会大大增加意外冲突的风险。为此，美国应该：（1)明确和公开确认美国现有的政策，即只有人类才能授权使用核武器，并要求俄罗斯和中国做出类似承诺；（2）建立与竞争对手讨论人工智能对危机稳定性影响的渠道；（3）制定与人工智能和自主武器系统开发、测试和使用相对应的国际实践标准。

改革国家情报部门。由于情报部门将比任何其他国家安全机构都更加受益于人工智能带来的好处，因此它们应在整个工作流程中，包括情报收集和分析在内，全面采用和结合人工智能。为了充分利用人工智能，国家情报总监办公室需要向旗下的科技领域负责人进行授权并提供资金和设备等各种资源。整个情报系统应在其分析中利用开源和公开信息，确定科技情报收集的优先级。为了更好地获得洞见，情报机构需要开发人机编组创新方法，通过人工智能增强人类判断。

在政府中增加数字人才。目前，国家安全机构需要更多的数字专家来弥补自身在采购、建造和使用人工智能及相关技术方面的不足。国防部和情报部门的人才短缺将是到2025年实现人工智能战备状态的最大障碍。政府需要建立新的人才管道，包括组建美国数字服务学院来培训现有和未来的员工；政府需要组建一支国家数字人才预备役部队，招募包括行业专家、学者和大学应届毕业生在内的各类民间人才；政府需要按照陆军医疗团模式组建数字军团来组织现役技术人员。

建立对人工智能系统的合理信心。如果人工智能系统未能按照设计目的进行工作，或者工作表现偏离预期并产生重大负面影响，在领导人层面，人工智能不会获得批准；在运营商层面，人工智能不会获得应用和推广；在国会层面，人工智能不会获得资助；在民众层面，人工智能不会获得支持。因此，为了建立合理信心，政府应着重确保人工智能系统的耐用性和可靠性，包括对人工智能安全研究进行投资，并通过由国家级实验室主导的持续性研发活动推进人机编组。随着人工智能系统在数量、范围和复杂性上的增加，还应增强国防部的测试和评估能力。政府应任命人工智能事务高级负责人来改善行政领导和政策监督。

制定人工智能国家安全用途的民主模式。人工智能工具对美国情报、国土安全和执法机构至关重要。公众对人工智能的信任取决于合理的保证，即政府在使用人工智能时，应尊重个人隐私、公民自由和公民权利。政府必须赢得这种信任，确保合法有效的使用人工智能工具。为此，政府需要开发人工智能工具来监督和审计人工智能应用，提高人工智能应用的公众透明度；政府需要确保受其人工智能行动影响的团体和个人能够获得补偿，并且建立相应的流程；政府需要加强监督和治理机制，建立特别工作组来处理公众对人工智能和个人隐私、公民自由和公民权利日益增长的关切。

第二部分：赢得技术竞争

目前，世界各国在人工智能和相关技术研究、开发和部署领域的竞赛正在加剧技术竞争，而这种技术竞争为更广泛的战略竞争提供支持。中国有完善的组织、丰富的资源、坚定的决心赢得这场竞赛。美国在关键领域仍然保持优势，但目前的趋势令人担忧。鉴于学术界和商业界在深层次上的互通互联变得日益复杂，我们无法针对这场竞赛给出强有力的回应，但是美国必须尽一切努力保持其在世界范围内的创新领导地位。美国政府必须拥抱人工智能竞争，并通过组织协调和资源调整来赢得竞争。

在白宫领导的技术竞争策略指导下开展组织活动。美国必须将人工智能的考量因素从技术层面上升到战略层面。目前，由人工智能主导的新兴技术支撑着美国的经济繁荣、国家安全和人民福祉。因此，白宫应建立一个由副总统领导的技术竞争力委员会，从安全、经济和科技角度对人工智能的各种考量因素进行整合；制定全面的技术战略并监督其实施。

赢得全球人才竞争。如果美国无法在国内培养更多的人才，也无法从国外招募和留住更多的现有人才，那么我们就有可能彻底输掉这场全球性的人工智能竞赛。美国必须在两大方面采取积极行动：一方面，国会应该通过《国防教育法案》（第二版）来解决当前教育体系的缺陷，包括在 K-12 教育和就业技能再培训中增设人工智能课程，向在人工智能关键性领域从事学习和研究的本科生和研究生提供奖学金；另一方面，国会应该实施全面的高技能移民政策，通过新的激励措施和签证、绿卡和工作可移动性改革，鼓励更多的人工智能人才在美国学习、工作和定居。

加快国内的人工智能创新。政府必须对人工智能研发进行重大投资，并建立全国人工智能研究基础设施，向平民提供更多的人工智能研发资源，在民间普及人工智能研发活动，进而推动人工智能发展。政府应：（1）到 2026 年，每年为人工智能研发提供双倍的非防卫资金（320 亿美元），建立国家技术基金会，使人工智能研究所的数量增加三倍；（2）通过建立由云计算资源、测试台、大型开放培训数据和开放性知识网络组成的全国人工智能研究基础设施，扩大接触人工智能的机会，并为科学和工程新领域内

的各项实验提供支持；(3）通过创建人工智能市场和形成区域性创新集群，加强商业竞争力。

执行全面的知识产权政策和制度。美国必须明确知识产权政策是保卫美国在人工智能和新兴技术领域领导地位的国家安全优先事项。鉴于其他国家常常以不正当手段窃取和使用美国的知识产权，这一点尤其重要。美国缺乏人工智能时代所需要的全面知识产权政策，并常常受到现行专利资格和专利性原则中法律不确定性的阻碍。美国政府需要从国家安全优先事项的角度重新规划知识产权政策和制度，推动知识产权相关的一系列改革。

建立微电子产业设计和制造的国内基础。在领先全球微电子行业数十年之后，今天美国几乎完全依赖其他国家生产尖端半导体，为对防御系统至关重要的人工智能算法提供动力。简单地说，如果政府不采取一致行动，国内的先进芯片供应链将面临风险。尽管重建国内芯片制造产业代价昂贵，但是我们仍然不得不立刻采取行动。美国应该坚持在最先进的微电子领域保持至少领先中国两代的战略，并通过提供资金和奖励等手段，激励国内尖端微电子制造业开展多样化生产。

保护美国的技术优势。随着美国的技术优势逐渐消减以及其他国家企图获取美国技术情报和两用技术的活动不断增加，美国必须重新审视如何在不过度阻碍创新的前提下，保护创新理念、技术和企业的完美解决之道。美国必须：

首先，为了更好保护关键性的两用技术，实现出口管制和外国投资筛选现代化，具体措施包括：建立监管部门，全面推动近期立法改革，与盟国协调对先进半导体制造设备的出口管制，扩大对来自竞争对手的海外投资者的披露要求。

其次，将国内科研型企业作为国家资产加以保护，向政府机构、执法机构和研究机构提供工具和资源进行细致入微的风险评估，分享有关特定威胁和战术信息，与盟友和合作伙伴协调研究保护工作，加强对研究机构网络安全支持，通过加强签证审查限制可疑研究合作。

构建有利的国际技术秩序。美国必须与盟国和合作伙伴一道开展工作，通过促进新兴技术应用来强化民主规范和价值观，通过协调政策和投资来推动数字基础设施和技术在全球范围内的应用，捍卫国际技术标准的完整性，合作推进人工智能创新，通过分享实战经验和资源来防止技术的恶意使用和专制政权在民主社会的影响。美国应该建立并领导新兴技术联盟，从而实现上述目的。与此同时，美国还应建立多边的人工智能研究所，提高美国作为新兴技术全球研究中心的地位。美国国务院应该进行重新定位、重组并获得更多资源，从而有效领导新兴技术领域的外交工作。

赢得技术竞赛。美国有必要争取和维持人工智能领域的领导地位，但是这远远不够。今天，人工智能已经成为新兴技术的核心，推动其他技术的发展，接受其他技术的反哺。因此，美国必须制定一份权威的技术清单，详细罗列能够支撑美国在 21 世纪仍然保持强大竞争力的各项技术，包括人工智能、微电子、生物技术、量子计算、5G、机器人技术和自主系统、添加剂制造和能源存储技术。为了确立在上述领域的领导地位，美国需要在特定平台进行投资，推动实现变革性突破，并在平台中建立充满活力的国内制造生态系统。与此同时，政府应当对即将出现的新兴技术进行持续识别和优先处理。

结论

这个全新的竞争时代将改变人类赖以生存的世界和长久以来的生活方式。在这场伟大的变革中，我们可以选择勇立潮头，也可以选择随波逐流。我们知道，人工智能在生活的方方面面将会继续得到普及，而人工智能技术创新的步伐将会继续加速；我们知道，美国的竞争对手意志坚定，它们积极打造基于人工智能的全新军事能力与我们进行对抗；我们知道，中国政府和人民下定决心在人工智能领域实现对美国的反超，夺得人工智能全球霸主的地位；我们知道，人工智能技术的进步与发展取决于这项技术本身，它能够给我们带来先发优势。现在，我们必须采取行动。我们建立的原则、进行的联邦投资、部署的国家安全应用、重新设计的组织结构、建立的合作伙伴关系、组建的各种联盟以及培养的各种人才将为美国战略铺平道路。美国应积极投入各种资源来维持人工智能创新的领导地位，以负责任的态度利用人工智能来捍卫自由的人民和自由的社会，并为全人类的福祉推进科技的进步。人工智能将重组世界，美国必须引领变革。

序言

人工智能国家安全委员会向总统和国会提供旨在“推动人工智能、机器学习和相关技术发展，以便全面解决国家安全和防务需求”的建议。根据《约翰·s·麦凯恩 2019 财年国防授权法案》第 1,051 小节的有关规定，委员会对人工智能开展审查。在进行此类审查时, 委员会应考虑：国家竞争力；保持技术优势的手段和方法；国际合作和竞争力的发展和趋势；在基础研究和高技术研发中，强化重点和投资手段；劳动力与培训；在军事上应用人工智能的潜在风险；伦理考虑；建立数据标准，鼓励数据共享；人工智能的未来演变[1]。

15 名委员分别由国会及行政部门进行提名。他们来自各行各业，成员包括技术人员、企业高管、学界领袖和国家安全研究人员。在国会两党合作框架下，他们出色完成了全部审查，并就最终报告达成共识。一直以来，委员会在“行动需求”和“透明度的重要性”两大原则指导下进行运作。

行动

委员会的工作成果主要包括：初次报告（2019 年 7 月），中期报告（2019 年 11 月和 2020 年 10 月），两份季度备忘录，关于新冠肺炎的一系列论文，以及最终报告。2019 年春季，当委员会正式完成组建并开始运行时，在最终报告中给出建议并不是我们的初衷。在一个独立的时间点上，评估人工智能等动态技术对国家安全的广泛影响就像带着镣铐跳舞一样困难。与此同时，科学家们持续实现人工智能技术的新突破，商业精英们正在以惊人的加速度寻找应用人工智能的新方法，全球范围内的竞争对手正在积极制定人工智能战略并为此投入大量资源。在这种背景下，委员会持续给出建议，希望跟上人工智能发展的脚步并为国会和行政部门的相关决策提供有参考价值的信息。目前，国会在《威廉·M（Mac）·索恩伯里 2021 财年国防授权法案》中采取了我们提出的诸多建议[2]，而行政部门同样也将我们的建议纳入其发布的政策和指令当中。我们一直向广

大的利益相关者学习并向其提供相关的教育，希望美国民众能够在人工智能对国家安全的影响问题上达成共识。

透明度

人工智能国家安全委员会始终坚持透明原则。作为一家联邦顾问委员会，我们先后召开了五次公开性的全体大会，进行了总计近 15 个小时的审议，对会议全程进行了在线直播，并将会议记录上传至委员会的官方网站。委员会对基于《信息自由法案》提出的 24 项申请进行了回应，发布了超过 2,500 页的材料，并在官网上公布了 700 页的意见稿接受公众的审查和评价。部分材料由于涉及国家安全机密无法公开，除此之外，委员会积极采取各种措施确保工作的透明度。在每次全体会议结束及发布季度报告和向国会提交书面意见之后，我们主动与媒体进行联系，邀请他们进行采访并召开新闻发布会。在数十次新闻发布会上，我们分别与非政府组织、联邦政府组织和国际组织进行合作，共同向媒体和公众披露我们提出的各项建议。

最终报告

委员会在最终报告中给出的建议可以作为美国制霸人工智能时代的行动指南。报告主体共分为 16 章，每一章都包含了参考价值极高的若干建议。此外，我们特别在报告主体后增加了“行动蓝图”部分，概述了政府部门和机构执行委员会建议时应采取的具体步骤。值得注意的是，我们对这些步骤尽可能地进行了详尽的描述，包括提供立法文本和行政命令草案，以便帮助总统和国会迅速从理解人工智能转向采取具体行动，从而造福美国人民。

最终报告代表的发布是委员会的一个里程碑，但它绝不会是我们的最后一项行动。在剩余任期内，委员会将重点关注执行情况，协助总统和国会进行投资和行动，进而赢得人工智能时代。

---

[1] 参见 Pub. L. 115 -232, 132 Stat. 1636 (2018), 请访问 https://www.congress.gov/115/bills/hr5515/ BILLS-115hr5515enr.pdf 获取全文。

[2] 参见 Pub. L. 116 -283, 134 Stat. 3388 (2021), 请访问 https://www.congress.gov/bill/116th-congress/ house-bill/6395/text 获取全文。

目录

引言

人工智能技术是近几代以来最强大的工具，它可以丰富知识、促进繁荣、增加人类经验。对于掌握人工智能技术的国家来说，它将是巨大的力量源泉。人工智能将促使各国政府和企业之间在人工智能领域展开竞争。与此同时，民族国家还将利用人工智能来实现战略雄心。

目前，美国民众尚未充分认识到人工智能革命对社会、经济和国家安全的深远影响。近期出现的人工智能突破（例如一台计算机在流行的策略游戏围棋中击败了一位人类选手[1]）震惊了全球，世界各国纷纷采取有针对性的行动。与其他国家的摩拳擦掌不同，美国对此无动于衷。尽管我们的私人部门和大学在人工智能领域占据全球领导地位，但是美国对于人工智能时代的带来没有进行任何准备。美国民众必须意识到，如果美国想在这场创新竞赛中获胜的话，政府必须发挥坚定的领导作用。国会和政府必须为实现这一目标提供所需的公共资源支持。

这是一个技术机会井喷和美国战略颓势凸显相重叠的年代。中国是一个强有力的竞争对手，拥有挑战美国技术优势、军事优势和全球诸多领域内霸主地位的实力、人才和雄心。与此同时，人工智能正在加深网络攻击和虚假信息活动带来的威胁，俄罗斯、其他国家和非国家行为体正在利用这些攻击和虚假信息活动渗透我们的社会，窃取我们的数据，干预我们的民主，而这不过是冰山一角。同时，新冠肺炎肆虐和气候变化所体现的全球危机，凸显了扩大国家安全概念和寻找创新技术解决方案的必要性。人工智能可以帮助我们有效应对这些新挑战。

我们是幸运的一代。人工智能革命并不是战略意义上的“秘密武器”。今天，我们在日常生活中，无时无刻不在体验人工智能带来的影响。因此，在我们不得不面对人工智能对国家安全所有领域的影响之前，事实上我们能够预见到相关研究进展如何转化为现实世界的具体应用。委员会对人工智能给国家安全带来的挑战发出警告，同时对其正面影响进行评估，而不对此前忽视的警告和错失的机会进行解释。目前来说，我们仍然拥有进行变革、建设更加安全的世界和开创更加美好未来的“窗口期”。人工智能创新的步伐并非一成不变，而是以加速度飞快向前。如果美国不采取行动，在下个十

年内，它很可能将人工智能领导地位拱手让给中国，并且在面对许多国家和非国家行为体带来的人工智能威胁时，变得更加脆弱不堪。

委员会认为，在人工智能时代，美国需要执行特定的防御和竞争策略。白宫必须主导一系列政府重组和国家重新定位行动。最终报告给出了上述策略的核心元素。

- 第一部分“在人工智能时代保卫美国”（第 1-8 章）概述了美国必须采取哪些行动抵御来自国家和非国家行为体的一系列人工智能相关的威胁，并就美国政府如何以负责任的态度利用人工智能技术来保护美国人民及国家利益给出建议。

- 第二部分“赢得技术竞争”（第 9-16 章）概述了人工智能在更加广泛的技术竞争中的作用。每个章节分别论述了这场竞争的关键性元素， 并就美国政府必须采取哪些行动推动人工智能创新、提高国家竞争力和保护美国的关键性优势给出建议。

人工智能关系重大的原因

1901 年，托马斯·爱迪生预测了电力对人类的影响。在电灯泡问世 20 年之后，他前瞻性地预见了一项带有无限可能的通用性技术。“（电力）是所有领域的核心……它拥有改变人类生活的秘密。[2]”人工智能是一项非常不同的通用技术，但是和爱迪生一样，今天我们也正站在相似的历史节点上，并且能够看到同样广泛的影响[3]。计算机系统解决问题和执行任务的能力正在飞速提高，在某些情况下甚至超过了人类的表现，这在很多情况下改变了人类生活和科学领域。人工智能将融入到几乎所有的未来科技当中，而支撑美国经济发展和国家安全的创新基础也将受到人工智能的影响。作为“所有领域的核心”的人工智能，它的善意或恶意用途，将重组这个世界。

委员会的评估根植于对人工智能当前发展状态的理解和对这项技术发展趋势的预测。

*今天，人工智能已经渗透到我们日常生活的方方面面，创新的步伐也在不断加快。*人工智能从大大小小的方面改变了我们的生活，对此我们已经习以为常。举例来说，一部“智能”手机集成了多项人工智能技术，包括语音助手、照片标签、面部识别、搜索应用、推送和广告引擎以及操作系统中隐藏的人工智能强化程序。目前，人工智能正在协助人类对流行性疾病的爆发、蔓延和升级进行预测，设计和优化产品和服务配送，监督交通流量和交通安全，加速药品和医疗发明，以及实现办公自动化。认识到变革的速度有助于我们理解人工智能的威力。在科技攻关活动中应用人工智能技术，可以大大缩短创新所用的时间，并将多个学科领域的科学构想变成现实。

部署和应用人工智能仍然是一个难题。人工智能无法魔法般地解决问题。随着人工智能从精英学科转变为主流工具，工程学突破将与科技突破一样重要。不同行业内的人工智能早期应用者已经认识到：即使人工智能技术已经成熟，但是部署人工智能仍然困难重重。探索理解现有人工智能算法和系统（部分算法和系统存在的时间已经长达十多年）的实际应用将给现实世界带来很多重大影响。人工智能技术与现实世界的融合面临巨大挑战。利用数据、强化和打包实验室算法，使其可以应用于具体领域，在老旧设备和僵化组织中部署人工智能软件，都需要时间、行动和耐心。人工智能的整合融入通常需要克服大量的组织和文化障碍，以及自上而下的领导和推动。

人工智能工具普及范围加大，普及速度加快。先进的深度学习技术由于价格昂贵，常常让人望而却步，并且需要海量数据、强大的计算能力和专业知识。尽管如此，人工智能不再只是少数几个大国或大型企业的专利。随着推动人工智能应用的多个机器学习工具已经面向大众公开并可以被非专业人士使用，开源应用程序和开发工具与廉价的云计算和数据密集度低的方法相结合，人工智能带来的重大发展机遇已经扩展到全球，包括国家和非国家行为体。

人工智能正在改变人机关系。在现代社会中，我们对机器和自动化技术的依赖程度或许已经超出了想象。举例来说，在过去的几十年内，美国军方一直使用自主系统。尽管如此，随着人工智能能力的提升，人机“团队”内部的动态将发生变化。过去，计算机只能按照人类清晰定义的参数或编程制定的一整套规则去执行任务。随着人工智能能力的提升，计算机可以在人类未能清晰定义参数的基础上进行学习和执行任务，以前所未有的体量和速度创建选项并采取行动。在人类活动的许多领域内，人工智能创新也带来了许多重要问题，例如向智能机器授权的行动选项、情形和原因。在国家安全领域，由于防御系统和情报系统与人工智能之间的整合，这些问题显得尤为重要。因此，一方面我们需要正视人工智能的普及，一方面利用细致入微的方法、强烈的求知欲和谨慎的态度处理社会上出现的这种新的复杂性。

第一部分：在人工智能时代保卫美国

随着人工智能技术在社会其他方面的普及，它将对国际竞争和冲突产生同等影响[4]。我们必须利用人工智能来改变保卫美国的方式、震慑敌人、获取情报、英勇作战并获得最终胜利，同时确保担负保卫美国重任的军事人员和文职人员能够利用人工智能和相关技术，迅速安全地完成使命。

人工智能是一项典型的军民两用技术。利用人工智能，一台机器可以在复杂的环境内，以超过人类的准确率迅速进行感知、决策和行动。毫无疑问，这在任何领域内都代表巨大的竞争优势。因此，政府和非国家团体将在军事领域应用人工智能。

我们希望对人工智能能力进行大规模普及和推广。多款人工智能国家安全应用只需要少量资源和适当的专业知识就可以使用。人工智能算法通常可以轻易获取，硬件是“现成的”并且在大多数情况下，可以向个人公开出售（例如图像处理单元等）。任何人都可以轻松下载和使用换脸工具“深度伪造”[5]，人工智能使能工具和变异的恶意软件掌握在黑客手里[6]，而廉价的致命性无人机也将越来越常见。2020 年 10 月，阿塞拜疆在与亚美尼亚的战争中使用了土耳其的无人机和以色列的巡飞弹，这证明了自主型军事能力正在普及[7]。许多国家对此保持密切关注并且从中吸取了经验。因此，流氓国家、罪犯或恐怖分子不计后果或不道德地使用人工智能技术的可能性正在增加。

在新时代新冲突的背景下，人工智能能力将成为首选工具。对于那些决心挑战美国而又避免直接军事对抗的国家和非国家行为体而言，它们可以利用人工智能来强化现有工具并开发新的工具。目前，对手利用我们的数字开放性，通过人工智能，发动信息战和网络攻击。多年来，广告公司和技术公司已经充分意识到，机器学习是进行数据收集、分析和目标群体定位的强大工具。今天，我们的对手也意识到了这一点，因此广告技术将成为国家安全技术。通过间谍活动和公开数据，对手将收集大量信息并利用人工智能识别个人、社会和关键性基础设施的漏洞，从而模拟最佳操纵行为，并付诸于行动。

人工智能将全面改变军事事务。在人工智能的协助下，武装部队可以快速高效地进入战备状态、感知和解读外界信息、准确进行决策并采取有针对性的行动。许多武器系统将集成一种或多种人工智能技术。与此同时，人工智能系统将为指挥官提供更多选择，搭建连接所有领域的作战网络。它将改变物流、采购、培训及新硬件的设计与开发。采用人工智能需要开发全新战术和军事行动理念。未来，战争将成为算法之间的对抗。战场优势将从作战部队的规模和装备水平等传统因素，转变为高级数据收集和同化、连接性、计算能力、算法和系统安全性等因素。

竞争对手正在积极开发军事用途的人工智能概念和技术。俄罗斯计划实现大部分军事系统自动化[8]。目前，它极不负责任地在叙利亚部署了自主系统，并在战场上对其进行测试[9]。与此同时，中国将人工智能看作削弱美国常规军事优势的重要手段，大力推动新一代人工智能技术的“跳跃式发展”。中国军方对“智能化战争”持积极肯定的态度。举例来说，中国在无人机蜂群作战系统进行了投资，以此抗衡美国的海上霸权[10]。中国的军事领导人公开谈论利用人工智能系统进行“侦察、电磁对抗和协同火力打击[11]”。此外，中国正在围绕真实场景设计的军事游戏中进行算法测试与培新。当其他国家部署人工智能使能军事系统时，我们担心它们不会受到指导美国武装部队的严格测试标准和道德规范的约束。

人工智能将彻底改变情报实践活动。在国家安全领域中，最适合人工智能大展拳脚的功能板块莫过于情报获取与分析。智能机器将对从不同渠道收集的所有信息进行筛选，找出关键信息，执行翻译任务，整合来自不同领域的数据集，确定相关性和联系，并向分析师和决策者提交分析结果。加速推进人工智能在国家安全领域的应用最紧迫及最有说服力的原因在于，更加先进的机器分析可以在下次攻击开始之前找到四处分散的点并将其一一连接起来，而人类分析人员无法独立看清事情的全貌，因此也无法承担保卫美国人民安全和利益的重任。

赤手空拳对阵具有人工智能能力的对手，将打开灾难的大门。人工智能可以将决策时间框架从分钟压缩到秒，扩大攻击规模。与此同时，它要求相关人员及时做出反应，而这往往超出了人类的认知极限。因此，在缺少人工智能使能机器协助的情况下，人类作战人员将无法抵御人工智能支持的网络或虚假情报攻击、无人机蜂群或导弹攻击。即使最出色的人类作战人员也无法抵御由每秒进行数千次机动，并以高超音速移动的多台机器在人工智能技术支持下进行的多域作战。人类本身无法同时覆盖所有战场，但是智能软件可以做到这一点。

以令人信服的逻辑、较快的速度和谨慎负责任的态度部署人工智能应用。政府在部署应用人工智能时，应该遵循传奇篮球教练约翰·伍登的原则：要快，但不要急[12]。与其他“安全关键的”人工智能应用一样，军事和情报领域的人工智能在开发和部署之前需要慎重考虑。目前，部分人工智能系统在实际应用过程中，常常暴露出缺少弹性和容易瘫痪的漏洞。因此，必须对应用于军事和情报领域的所有人工智能进行严格和测试并采取充分的保障措施，同时深入了解人工智能在现实世界中的具体应用和测试台上的表现存在哪些差异。人工智能支持的自主武器系统可以更加精确地打击目标，从而减少无意中造成的平民伤亡。但是，它们也带来了重要的伦理问题，即人类在判断是否使用致命武力时起到的作用。如果设计或使用不当，这类武器也可能增加军事升级的风险。

对于以负责任的态度在国防和情报领域使用人工智能的原则，社会各界正在形成共识[13]。人工智能赋能的机器在工作时，如果表现未能达到设计预期或遵循清晰的原则，则作战人员、组织团体和美国民众将不会使用、接受或支持人工智能。因此，仓促推广和普及人工智能是一种危险的举动并可能导致不良后果。我们必须小心谨慎，以免美国民众对于人工智能带来的好处失去信心。与此同时，我们必须清楚地认识到，风险不可避免。未能利用人工智能解决当前面临的国家安全挑战将使美国处于不利地位，削弱美国军人的作战能力，并且将纳税人的钱浪费在过时低效设备上的优势；推迟人

工智能的普及和应用将把所有的风险转嫁给我们的下一代——他们很可能不得不使用 20 世纪的工具来保卫 21 世纪的美国，甚至与 21 世纪的对手交战。

> “即使最出色的人类作战人员也无法抵御由每秒进行数千次机动，并以高超音速移动的多台机器在人工智能技术支持下进行的多域作战。人类本身无法同时覆盖所有战场，但是智能软件可以做到这一点。”

美国政府仍然以人类而不是机器的速度进行运作。采用人工智能需要从战术到战略层面、从战场到五角大楼，对国家安全业务实践、组织文化和思想状态上进行深刻调整。目前，在包括基础业务自动化在内的大多数人工智能领域，政府都远远落后于已经采用最先进人工智能技术的商业界。而且，政府还受到了技术赤字的严重影响，具体表现为数字人才短缺、采购政策不到位、网络架构不充分和数据实践表现差等。与此同时，官僚作风也阻碍了政府与人工智能技术领袖之间建立深层次的合作伙伴关系。政府必须实现角色转变，成为更受欢迎的采购商与合作伙伴。在缺少重大战争和恐怖袭击等外界因素的刺激下，国家安全领域内的创新，需要强有力的领导力。

第二部分：赢得技术竞争

除了在国家安全和防务领域内的有限应用以外，人工智能已经成为全球技术竞争的支点。它将从各个维度推动国家实力的增长，并对医疗、食品生产和环境可持续发展等行业和领域进行赋能。在相近领域和技术范畴内成功采用人工智能将驱动经济发展，塑造理想社会，决定哪些国家可以在全球范围内施加影响和扩展势力。尽管很多国家已经制定和出台了本国的人工智能发展战略，但是只有美国和中国具备领导全球人工智能产业发展所需的资源、商业实力、人才库和创新生态系统。在部分研究和应

用领域，中国已经成长为与美国旗鼓相当的对手；甚至在某些应用领域，中国的技术更加先进[14]。在未来的十年内，中国将超越美国，成为全球人工智能领域的超级大国[15]。

在公平的竞争环境下，美国有能力在创新领域超越任何对手。但是，由于美中两国的人工智能创新方法存在根本性的差异，今天美国的人工智能领导地位岌岌可危。数十年来，美国的创新模式一直是其他国家羡慕的对象：开放的思想交流、自由市场和在支持基础研究时的有限政府参与构成了美国创新方式的支柱，代表了美国的价值观。在美国，技术企业通过竞争的方式获得市场份额，它们不是国家权力的工具。研究人员在开放的研究环境中进行合作，与处于同一水平线的对手展开竞争，致力于实现人工智能的突破，而无需考虑合作伙伴的国籍问题。企业竞相寻求利润和推出下一个重大理念，也为风险资本的国际流通和人工智能相关的贸易活动注入了活力。

在美国，大多数人工智能领域的进展由私人部门和大学完成。我们不能失去自下而上的创新文化和车库起家的创新精神。但是，在这场战略竞争中，完全分布式的研究方法并不是致胜之道。即使大型的科技企业也无法与中国的资源进行抗衡或进行使美国保持技术领先的巨额投资。我们需要将政府和私人部门的研究活动有机的结合起来，以便赢得这场技术竞赛的胜利。

中国有完善的组织、丰富的资源、坚定的决心赢得这场竞赛。人工智能对于中国的全球扩张、经济和军事实力和社会稳定至关重要。目前，人工智能发展规划已经成为中国在若干关键和新兴技术领域领先世界的总体战略的一部分。更重要的是，中国在执行人工智能发展规划方面，已经领先于其他国家。2017 年，中国提出了到 2030 年在新一代人工智能领域建立世界领先地位的发展战略、总体目标、分阶段目标和具体时间表，信心坚定的中国领导人承诺为实现这一目标提供资源支持[16]。目前，在中央政府指导下，通过人才招聘、技术转移和投资等手段，中国正在系统性地实施海外人工智能知识成果吸收和转化工程。它还出台了雄心勃勃的人才培养计划，建立新型人工智能中心，旨在打造和培养新一代的人工智能工程师，大力扶持龙头企业（包括华为、百度、阿里、腾讯、科大讯飞和商汤科技）引领国内的人工智能技术发展，推进以军民融合为主题的国家重点军事和安全项目，占领海外市场[17]，同时向全球范围内的大量数字基础设施建设项目提供资金。中国出台了知识产权战略，并且正在制定全球性的人工智能开发技术标准[18]。而且，根据中国相关法律的规定，在特定情况下，企业有义务向上级主管部门公开数据[19]。

人工智能进步推动了包括电子商务、搜索引擎和社交媒体在内的广泛的平台技术竞争。在计算、数据、人才和商业化方面赢得人工智能竞赛的国家、企业和研究人员将着眼于赢得更大的胜利。从本质上来说，消费者/参与者的基数越大，数据的数量就

越大，质量就越高，算法就越优秀，进而带来更多的用户和数据以及更出色的性能。最终，少数几家企业将会成为占据主导地位的平台。如果中国企业在这些竞赛中胜出，它不仅仅会使美国企业处于不利地位，同时也会为美国及其盟友即将面临的地缘政治挑战打下坚实的基础。在海外获得平台主导权使得中国可以获取用户的数据，并将国内控制体系延伸至海外。一旦中国掌控了数字基础设施、社交媒体和电子商务平台，它都将拥有更大的影响力和实力来传递自己的声音，并塑造具有中国特色的新世界。

美中两国的人工智能竞赛由于两国之间的深度互联，变得异常复杂。上个世纪，美国和前苏联进行太空竞赛时，双方在平行的轨道上各自运行，两国企业独立进行研发活动，彼此之间的商业往来极少。今天，美国和中国的这场人工智能竞赛呈现出完全不同的特质。研究项目共享、人才流通以及包括供应链、市场与合资型研究企业在内的商业联系，使得两国的研究生态系统紧密相连。一直以来，美国进行的基础性研究及美国企业都从与中国的技术联系中获得了巨大收益；因此，一旦切断这种联系，将会造成适得其反的效果。然而，美国必须保护开放式研究的完整性及本国企业的知识产权，并通过出口管制和投资筛选等手段，保护对于美国国家安全至关重要的技术产业。

美国在关键领域保持优势，但是趋势令人担忧。与过去相比，今天世界上最优秀的科技人才更愿意在国内工作或移民海外[20]。美国在微电子（所有人工智能赖以运行的硬件系统）领域的领先优势逐渐被削弱，而向美国提供先进芯片的亚洲供应链和制造商又容易受到国内外势力的干扰[21]。目前，尽管许多机器学习工具已经得到普及而且单位计算成本已经下降，但实现尖端的深度学习研究突破所需的计算能力和数据访问使得象牙塔内的研究人员和小型企业很难就此展开竞争[22]。创新布局仍然集中在美国的少数几个区域[23]。

美国政府必须采取实际行动提高国家技术竞争力。政府有责任促进多样化和有弹性的研发生态系统和商业体系的发展，扩大人才渠道以吸引世界上最优秀的人才，加大人工智能领域本土人才的培养力度；政府有义务审慎而积极地保护关键的人工智能知识产权和打击非法的知识转移行为，通过立法手段和联邦奖励政策保护美国的硬件优势和构建供应链弹性；政府应实施以美国为主导的外交政策，与志同道合的盟友和合作伙伴共同建立一个国际联盟，确保人工智能的民主愿景，塑造数字未来。

人工智能竞赛需要白宫的领导。美国人工智能发展战略的关键元素涉及国家安全、经济和技术政策，由于过分复杂，任何一个部门或机构都无法担负起领导责任。只有来自白宫的强有力的行政领导才能推动政策运行，实现各方势力的平衡，动员全国进行必要的投资。

人工智能的用途、相关技术和价值观

世界各国政府广泛采用人工智能，这不仅影响着各国之间的国际秩序，也影响着各国内部的政治秩序。人工智能未来的利害关系与独裁和民主政治制度及意识形态之间的长久对峙密切相关。

技术本身并不具有任何意识形态，但是技术的开发方式、用途和适用的监管法律则反映了技术开发者和使用者的侧重点和价值观。不久，大多数政府都将具备人工智能监视和分析能力。随着技术的普及和扩散，国家之间的主要差异不再是技术的质量或复杂性，而是表现为技术的使用目的、使用方式和适用规则。

独裁政权将继续利用集成了人工智能技术的人脸识别、生物特征识别、预测性分析和数据融合作为监视、影响和控制的工具。目前，部分国家利用人工智能监视技术压迫少数族裔和监督公民言行，这意味着独裁政权将利用人工智能系统来巩固独裁统治、追踪本国公民的踪迹和数字行为以及压制异见[24]。数字系统的全球性流行为在更大范围内实行独裁统治创造了条件。但是，自由的民主国家也在国内安全和公众安全领域使用人工智能。值得注意的是，超过半数的先进民主国家也在使用人工智能监视系统[25]，它们的人工智能技术用于合法的公共用途并且符合法治精神。不过，随着部分民主国家开始倾向于进行不民主的实践活动，以动摇和破坏依法治国根基的方式使用数字工具将导致民主的倒退。因此，我们必须持续对维护个人自由保持警惕。一个负责任的民主国家必须确保政府谨慎明智地使用人工智能，维护自由开放社会中标志性的个人权利和自由。

美国政府在开发和部署人工智能技术时，应当坚持充分的透明和强有力的监督，推行问责制度，避免恶意用途。美国政府仅仅表明自己对人工智能用于维护和巩固独裁统治的反对立场是远远不够的。我们必须向全世界证明，民主国家如何在坚持自由民主的价值观的同时，利用人工智能保护本国公民的安全。出于国家安全目的，我们亟需部署人工智能，打击在美国境内活动的外国和国内恐怖分子。此外，还需要确保人工智能的安全应用符合法律规定的个人自由和平等保护的核心价值观。

美国必须在民主国家联盟中发挥领导作用。在确保人工智能的开发和使用方式不会破坏国内的民主制度之后，我们还必须促进建立全球性的规范，使其他民主国家也可以安全地使用人工智能。尽管美国政府对其他国家治理实践的影响力十分有限，但是我们应当确保美国的外交政策议程中对于人工智能的使用进行了硬性规定：人工智能必须促进人权和对抗技术独裁趋势。美国可以通过外交手段，利用遍布全球的合作伙伴关系，倡导在国际机构中建立保护隐私的技术标准和规范；它还可以与志同道合的国家进行合作，确保其他国家除了接受中国的技术和社会治理方法以及获得能够保

护隐私等民主价值观的技术之外，还有其他选择。我们不寻求建立一个支离破碎的数字世界。相反，我们希望美国及其盟友生活在这样的一个世界里：在数字基础设施、电子商务和社交媒体方面有多种选择，不会受到独裁胁迫，享有言论自由、个人权利和隐私，并且能够容纳不同观点。

结论

新的竞争时代，新的起点。我们知道，人工智能在生活的方方面面将会继续得到普及，而人工智能技术创新的步伐将会继续加速；我们知道，美国的竞争对手意志坚定，它们积极打造基于人工智能的全新军事能力与我们进行对抗；我们知道，中国政府和人民下定决心在人工智能领域实现对美国的反超，夺得人工智能全球霸主的地位；我们知道，人工智能技术的进步与发展取决于这项技术本身，它能够给我们带来先发优势。现在，我们必须采取行动。我们建立的原则、进行的联邦投资、部署的国家安全应用、重新设计的组织结构、建立的合作伙伴关系、组建的各种联盟以及培养的各种人才将为美国战略铺平道路。美国应积极投入各种资源来维持人工智能创新的领导地位，以负责任的态度利用人工智能来捍卫自由的人民和自由的社会，并为全人类的福祉推进科技的进步。人工智能将重组世界，美国必须引领变革。

引言 – 尾注

[1] *Google DeepMind 挑战赛*，DeepMind（上次访问时间：2021 年 1 月 7 日），https://deepmind. com/alphago-Korea。

[2]援引自*奥里森*·斯韦特·*马登*（Orison Swett Marden）的《*他们是如何成功的：成功人士自述人生故事的生活故事*》，Lothrop 出版公司，238(1901)。

[3]吴恩达（Andrew Ng）进行这种比较时备受赞誉。请参阅例如纳尼·林奇（Shana Lynch），《*吴恩达：为什么说人工智能是新电力*》，《斯坦福商学院见解》（2017 年 3 月 11 日），https://www.gsb.stanford. edu/insights/andrew-ng-why-ai-new-electricity.

[4]有关人工智能和国际关系的概述，请参阅《德克萨斯国家安全评论》（2018 年 5 月）中迈克尔·霍罗威茨（Michael Horowitz）的《人工智能、国际竞争和均势》》，https://doi.org/10.15781/ T2639KP49。

[5]凯伦·郝（Karen Hao）和威尔·道格拉斯·海雯（Will Douglas Heaven）的 *Deepfakes 成为主流那一年*，《麻省理工学院技术评论》（12 月）https://www.technologyreview.com/2020/12/24/1015380/best-ai-deepfakes-of-2020/。

[6]尼古拉斯·杜兰（Nicholas Duran）等人，《*2018 年 Webroot 威胁报告*》，Webroot 2018 年），https://www-cdn.webroot.com/9315/2354/6488/2018-Webroot-Threat-Report_US-ONLINE.pdf；*人工智能对网络安全的影响：研讨会论文集*，国家科学、工程和医学学院（2019 年），https://www.nap.edu/catalog/25488/implications-of-artificial- intelligence-for-cybersecurity-proceedings-of-a-workshop；本·布坎南（Ben Buchanan）等人，*A 自动化网络攻击：*《炒作与现实》，安全与新兴技术中心（2020 年 11 月），https://cset.georgetown.edu/research/automating-cyber-attacks/；*深度利用：使用深度强化学习的全自动渗透测试工具*，GitHub（上次访问时间：2021 年 1 月 9 日），https://github.com/13o-BBR-bba/machine_learning_security/tree/master/Deepexplit。

[7]萝宾·迪克逊（Robyn Dixon）的*阿阿塞拜疆的无人机在纳戈尔诺-卡拉巴赫的战场上占有一席之地--展示了战争的未来*，《华盛顿邮报》（2020 年 11 月）https://www.washingtonpost.com/world/europe/nagorno- karabkah-drones-azerbaijan-aremenia/2020/11/11/441bcbd2-193d-11eb-8bda-814ca56e138b_story. html。

[8]瓦季姆·科久林（Vadim Kozyulin），《*人工智能军事化*》，斯坦利和平与安全中心（2019 年 7 月），https://stanleycenter.org/wp-content/uploads/2020/05/MilitarizationofAI-Russia.pdf。

[9]迪伦 马利亚索夫（Dylan Malyasov），*叙利亚的战斗测试揭示了俄罗斯无人驾驶小型坦克的缺陷*，国防博客（2018 年 6 月 18 日），https://defense-blog.com/news/army/combat-tests-syria-brought- light-deficiencies-russian-unmanned-mini-tank.html。

[10]艾尔莎 卡尼亚（Elsa Kania）在美中经济与安全审查委员会的证词，*关于技术、贸易和军民融合的听证会*（2019 年 6 月 7 日），https://www.uscc.gov/sites/default/ files/June%207%20Hearing_ Panel%201_ Elsa%20Kania_Chinese%20Military%20Innovation%20in%20 Artificial%20Intelligence_0.pdf；艾尔莎 卡尼亚，*“中国军事创新中的人工智能武器”*，布鲁金斯学会，1（2020 年 4 月 20），https://www.brookings.edu/wp-content/uploads/2020/04/FP_ 20200427_ai_ weapons_kania_v2.pdf。

[11]马库斯·克莱，*中国人民解放军的人工智能竞赛*，《外交家》（2020 年 11 月 5 日），https://thediplomat。com/2020/11/the-plas-ai-competitions/。

[12]安德鲁·希尔和约翰·伍登，*快速—但不要急于求成：*一生教诲中寻找成功之道，西蒙与舒斯特公司，69 (2001)。

[13]美国国防部新闻稿，*国防部通过了人工智能的伦理原则*（2020 年 2 月）https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod- adopts-ethical-principles-for-artificial-intelligence/；情报界的人工智能伦理原则，国家情报总监办公室（上次访问时间：2021 年 1 月 11 日），https://www.dni.gov/index.php/ features/2763-principles-of-artificial-intelligence-ethics-for-the-intelligence-community。

[14]格雷厄姆 • 艾利森（Graham Allison）和埃里克 • 施密特（Eric Schmidt），*中国是否正以人工智能的优势击败美国？*贝尔弗科学和国际事务中心（2020 年 8 月），https://www.belfercenter.org/publication/china-beating- us - ai - supremacy。

[15]请参阅亚历山德拉 • 穆萨维扎德（Alexandra Mousavizadeh）等人的*《全球人工智能指数》*，Tortoise 传媒（2019 年 12 月 3 日），https://www.tortoisemedia.com/2019/12/03/global-ai-index/。

[16]请参阅格雷厄姆 • 韦伯斯特（Graham Webster）等人，*全文翻译：中国‘新一代人工智能发展计划’*，新美国媒体（New America）（2017 年 8 月 1 日）https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/（翻译中国国务院关于印发《新一代人工智能发展规划》的通知，2017 年 7 月 20 日）。

[17]本杰明 • 拉森（Benjamin Larsen），*组建国家人工智能治理团队*，新美国媒体（2019 年 11 月 18 日），https://www.newamerica.org/cybersecurity-initiative/digichina/blog/drafting-chinas-national-ai-team-governance/（原文为格雷厄姆 • 韦伯斯特（Graham Webster）发表，*《人工智能政策和中国：国家主导发展的实际情况》*，斯坦福 - 新美国数码中国项目（2019 年 10 月 29 日）https://d1y8sb8igg2f8e.cloudfront.net/documents/DigiChina-AI-report-20191029.pdf）孟静，*中国将增强其“国家队”，以实现 2030 年全球人工智能领导地位的目标*，《南华早报》(2018 年 11 月 15 日），https://www.scmp.com/tech/innovation/article/2173345/china-boost-its-national-team-meet-goal- global-ai-leadership-2030；格雷戈里 • 阿兰（Gregory C. Allen），*了解中国的人工智能战略*，新美国安全中心（2019 年 2 月 6 日），https://www.cnas.org/publications/reports/understanding-chinas-ai - strategy（“SenseTime 和其他人工智能冠军公司被允许主导这些技术的代价是冠军公司与中国国家安全界的广泛合作。 即使在直接合作之外，中国在商业人工智能和半导体市场的成功也带来了资金、人才和规模经济，这既减少了中国因失去国际市场准入而带来的脆弱性，也为发展武器装备和间谍能力提供了有用的技术。”）

[18]布吕耶尔（Emily de La Bruyère）&皮卡西克（Nathan Picarsic），*中国标准 2035*、地平线咨询公司（2020 年 4 月），https://www.horizonadvisory.org/china-standards-2035 -first-report。

[19]默里 • 斯科特 • 坦纳（Murray Scot Tanner），北京的新国家情报法：*从防守到进攻*，“法律战计划”组织（2017 年 7 月 20 日），https://www.lawfareblog.com/beijings-new-national-intelligence-law-defense-offense。

[20]雷姆科 • 茨韦茨卢特（Remco Zwetsloot），*中国技术人才竞争方法：《政策、结果和正在形成的全球反应》*，布鲁金斯学会（2020 年 4 月），https://www.brookings.edu/wp- content/uploads/2020/04/FP_ 20200427_china_talent_policy_zwetsloot.pdf。

[21] “所有大批量、领先的[半导体]生产中的 90%将很快在台湾、中国和韩国进行。”国防部估计，到 2022 年，只有 8%的半导体将在美国进行制造，“相比 20 世界 90 年代的 40%，有所下降”。迈克拉 • 普拉策（Michaela D. Platzer）等人，*半导体：美国工业，全球竞争和联邦政策*，国会研究服务处，12（2020 年 10 月 26 日），https://fas.org/sgp/crs/misc/R46581.pdf（援引自里克 • 斯威策（Rick Switzer），微电子供应链集中在台湾、韩国和中华人民共和国对美国国家安全的影响，美国空军 [2019 年 9 月])。

[22]例如，非精英大学和人工智能初创公司难以负担训练复杂的机器学习（ML）模型所需的计算资源和数据费用。努尔 • 艾哈迈德（Nur Ahmed）和穆塔西尔 • 瓦赫德（Muntasir Wahed），人工智能的非民主化：*人工智能研究的深度学习和计算鸿沟*，阿奇夫论文预印本网站（2020 年 10 月 22 日），https://arxiv.org/abs/2010.15581.加强全国范围内的人工智能基础设施的必要性是人工智能促进协会发布的 20 年路线图的第一条建议。参见《美国人工智能研究 20 年社区路线图》，计算社区联盟和人工智能促进会，3，（2019 年 8 月)，https://cra.org/ ccc/wp-content/uploads/sites/2/2019/08/Community-Roadmap-for-AI-Research.pdf.

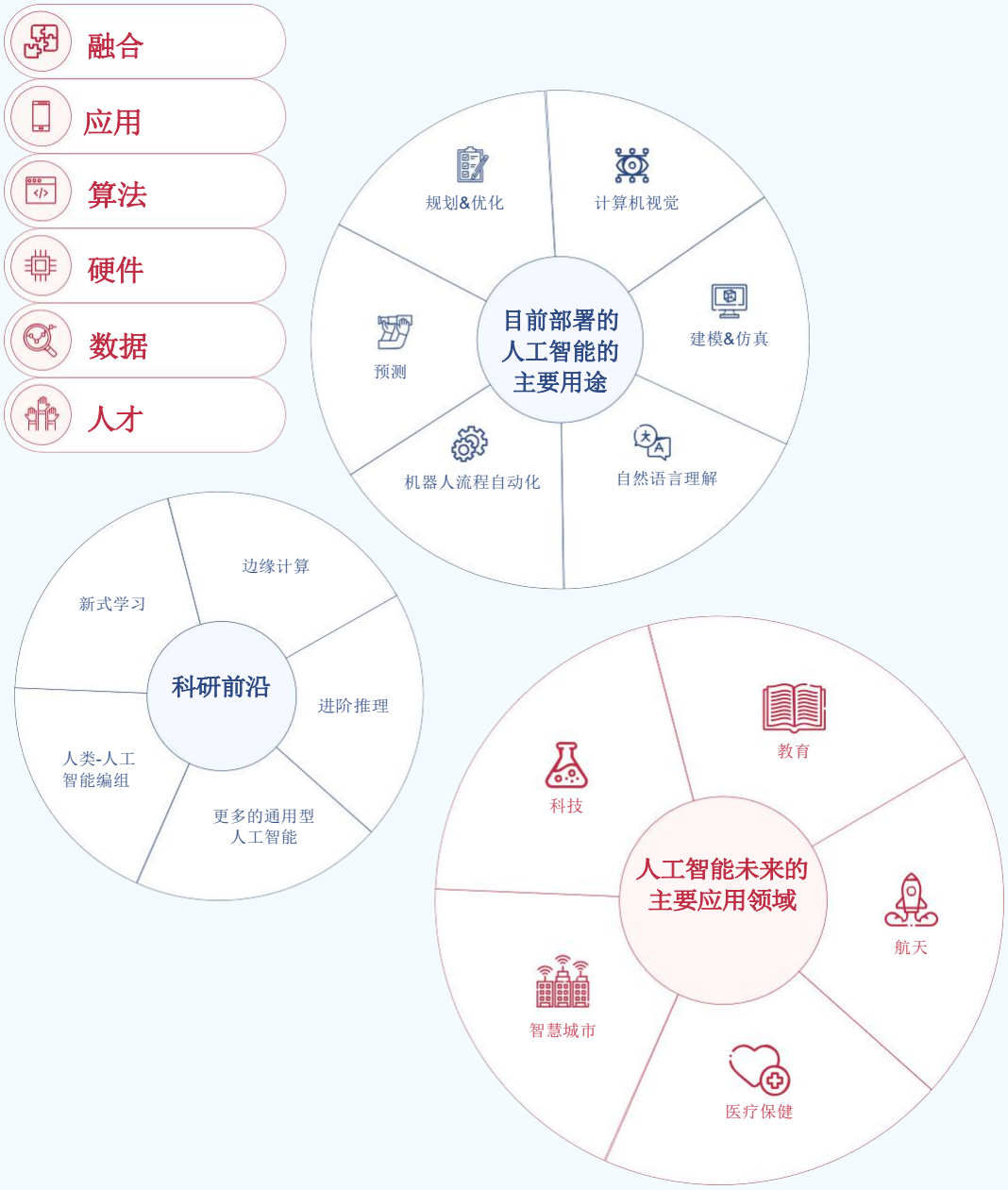
[23] 2005 年至 2017 年期间，美国创新部门 90%以上的就业机会仅发生在五个主要沿海城市。罗伯特 • 阿特金森（Robert D. Atkinson）等人，*增长中心案例：如何将科技创新推广到整个美国*，布鲁金斯学会（2019 年 12 月 9 日），https://www.brookings.edu/research/growth- centers-how-to-spread-tech-innovation-across-america/。

[24]参见帕特里斯 • 塔多尼奥（Patrice Taddonio），*中国政府如何对其维吾尔族穆斯林人口使用人工智能*，美国公共电视台前线频道（2019 年 11 月 21 日），https://www.pbs.org/wgbh/frontline/article/how-chinas-government-is-using-ai-on-its-uighur-muslim-population/；贝瑟妮 • 艾伦 • 易卜拉希米安（Bethany Allen-Ebrahimian），*曝光：我国大规模拘禁和逮捕算法操作手册*，国际调查记者同盟（2019 年 11 月 24 日），https://www.icij.org/investigations/china-cables/exposed-chinas- operating-manuals-for-mass-internment-and-arrest-by-algorithm/。

[25]参见史蒂芬 • 费尔德斯坦（Steven Feldstein），*人工智能监控的全球扩张*，卡内基国际和平研究院（2019 年 9 月），https://carnegieendowment.org/2019/09/17/global-expansion-of-ai- surveillance-pub-79847。

语境中的人工智能

人工智能不是单一的硬件或软件；相反，它是依赖于相关元素的技术组合，这些元素可以被看作堆栈。



人工智能不是单一的硬件或软件；相反，它是依赖于相关元素的技术组合。在处理如此宏大的课题时，委员会的立法授权对于研究工作的开展提供了指导意见，其中涉及两大类人工智能技术：第一，用于解决需要人类感知、认知、规划、学习、沟通或身体动作的任务；第二，以智能软件代理或实体机器人的形式，自主进行学习和行动[1]。

成功开发和部署人工智能技术取决于许多相互关联的元素，这些元素可以被看作堆栈[2]。人工智能需要人才、数据、硬件、算法、应用程序和集成。其中，人才是最基本的一项需求，它能够推动其他元素的创建和管理。数据对于大多数人工智能系统而言至关重要[3]。标记数据和策展数据使得目前的大部分机器学习可以用于创建新的应用程序及改进现有人工智能程序的性能。底层硬件为分析不断增长的数据池和运行应用程序提供计算能力。硬件层包括基于云的计算和存储，由网络通信主干网支持，用于连接网络边缘的智能传感器和设备。算法是通过数学运算，向系统下达数据传递指令，从而针对特定问题给出答案。应用程序使系统提供的答案有助于执行和完成特定任务。这些元素的集成对于部署端到端的人工智能系统至关重要。这需要大量的工程人才和投资来整合现有的数据流、决策管道、老旧设备和测试设计等。整合任务可能令人望而生畏，但在历史上却一直被低估[4]。

人工智能技术和应用，例如图案识别、机器学习、计算机视觉、自然语言理解和语音识别等，已经开展了几十年。在人工智能发展的早期阶段（这一阶段被 DARPA 描述为“第一波”），研究人员探索了许多方法，包括符号逻辑、专家系统与规划，并取得了一批富有成效的研究成果。部分研究成果建立在人类定义的“手工知识”基础之上，后来被用于机器推理和人机交互[5]。

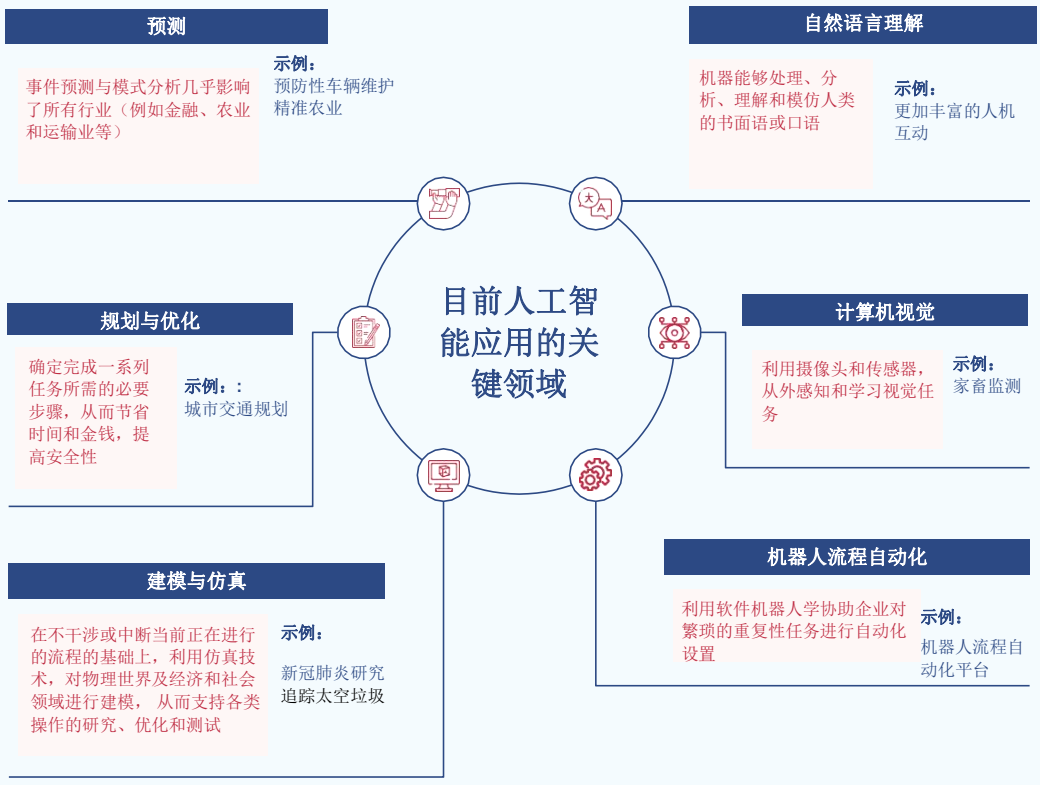
在过去的 10 年里，我们见证了由大规模统计机器学习推动的“第二波”人工智能浪潮，它使得工程师们可以创建模型。在给定样本数据或进行模拟交互的条件下，上述模型可以应用于特定的问题域。由于这些系统可以持续不断地从数据中进行学习，它们设计用来完成特定任务和实现特定目标。值得注意的是，这些系统表现出来的能力在某些方面类似于人类的认知过程：感知、推理、学习、交流、决策和行动。今天，我们部署的大多数大规模人工智能系统均采用了“第一波”和“第二波”人工智能方法。

部署人工智能技术的时代

今天，我们已经到达拐点：全球数字化转型带来海量数据供应；统计机器学习算法，特别是深度神经网络，尽管仍然存在局限性，但是已经能够熟练地解决问题[6]；支持机器学习的强大网络计算已经可以广泛地进行应用。基于上述因素，人工智能技术目前已经掌握在技术和非技术人员手中，由此引发了一个根本性问题——人们不再关注这项技术的工作原理，而是更加看重它能给我们带来的收益[7]。

目前，尽管人工智能技术仍然存在很大的局限性，但是它在某些特定领域的表现仍然十分抢眼。我们已经进入了部署人工智能的时代。人工智能的身影无处不在，它已经集成在智能电话、无线路由器和汽车等常用设备中,并且每天与我们进行互动。我们通常依赖人工智能技术含量较高的应用程序，例如通过移动电话或在线搜索新开业的饭店，设计出行方案，选择电影，获得客户服务等。

人工智能在关键领域的应用



我们很难预测人工智能的未来。五年前，很少有人能够预料近期在自然语言理解领域取得的重大突破，使得系统可以自动生成与人类散文几乎无法区分的全部文本[8]。过去五年，全球人工智能产业的投资显著增加[9]，普通研发活动的投资达到前所未有的水平，人工智能发展[10]并未出现放缓的迹象——相反，我们却看到了部署人工智能的新曙光。

“过去五年，全球人工智能产业的投资显著增加，普通研发活动的投资达到前所未有的水平，人工智能发展并未出现放缓的迹象——相反，我们却看到了部署人工智能的新曙光。”

人工智能技术的前沿

在下一个十年里，人工智能研究可能被定义为努力整合现有知识、推进新的学习方式，并使系统变得更加稳固、普及和值得信赖[11]。推进人机编组方面的研究以及混合型人工智能技术、增强型培训方法和可解释的人工智能方面的改进，将会成为学科前沿。

人机编组。掌握人类—人工智能协作与编组是人工智能未来应用的基础性要素。人类与人工智能之间的协同作用使得人机编组的效果呈指数级增长，大大超过二者的总和。目前，研究人员正在通过研究授权、可观察性、可预测性、可指导性和信任等问题来解决相关的挑战[12]。更深入地了解人类如何利用人工智能进行工作，将为我们创建有效的人类培训项目提供洞见，而积极推进语言解读领域的进步将为我们创建更加智能的系统——它能够总结复杂输入指令，并通过对下一代人机编组至关重要的类人对话进行交流。人机编组前沿领域包括需要在代理群组（人类和机器的混合编组或协调性机器人组成的团队）之间进行情报协作。

新的学习方式。新的学习方法能够提升培训及数据推断的效率[13]，降低对大量数据集的依赖，拓宽系统空间从而使得系统可以处理超出其原有工作范围之外的任务，建立情境式学习和常识推理通道。混合式人工智能技术通过整合不同的人工智能方法，能够充分发挥这些方法之间的互补性优势[14]。举例来说，神经-符号人工智能研究将符号处理与神经网络结合起来[15]。基于模型的方法和基于数据的方法之间也可以进行结合，例如在统计机器学习框架内利用物理知识[16]。研究人员还在积极推进利用少量标记数据支持监督式学习技术的发展[17]，而其他研究人员则已经设计完成了高效使用标记数据的方法[18]。通过模拟生成合成数据就是一种很有前途的方法[19]，它允许模型查看真实数据集可能没有遇到过的条件和场景，同时保留原始数据重要变量之间的相互关系及敏感数据的隐私[20]。

边缘计算。突破尺寸、重量和功率障碍同样也有助于人工智能的普及和隐私保护。目前，各大企业正在将更强大的计算能力装入紧凑型专用芯片内。这些芯片支持以更少的能耗培训和运行相同的模型，允许消费设备在本地运行复杂的模型，而不是在外部传输数据，然后等待模型远程运行。总之，将数据完全保留在正在训练或运行模型的设备上，可以有效保护人工智能驱动系统中的个人隐私[21]，代表着巨大的科技进步。

推理技术的进步。与人类相比，即使目前最出色的人工智能系统也缺乏“常识推理”能力，不过研究人员正在创建能够实现知识泛化及跨域学习转化的系统。人工智能系统在具备常识推理能力之后，可以模仿人类的能力对人和物体的物理属性、目的、意图和行为进行假设和分析，进而对行动或交互可能产生的后果进行描述。在分类和创建广义结构化本体及语言理解方面取得的进步将在机器理解语境和内容的同时，提升机器的学习能力，并且允许人类迅速发现问题的解决方案——过去，人们不得不花费数年的时间对问题进行审查[22]。这项研究有望为更可解释的人工智能和更强大的偏差检测和降低能力铺平道路，这对于提高更多的通用型人工智能技术的可信度至关重要[23]。

更加普遍的人工智能。迄今为止，人工智能解决方案已经展现了细致深刻的能力，但是与人类展示的能力之间有着根本的区别。人类在缺乏明确监督信号的情况下，通过学习执行任务；他们能够对完成一项任务所需的技巧举一反三，并将其用于执行其他任务；而且，利用常识性知识，他们可以积累经验、使用工具、进行推理。部分研究人员将“通用人工智能”作为实现人工智能超越小范围的垂直专业知识领域的目标。关于通用人工智能，争论的焦点主要集中在：是否会出现具体的突破，能够带来更为普遍的类人能力；是否在通用人工智能领域，研究人员将更有可能继续推动更通用的人工智能，沿着技能的一个或多个维度向前发展。不管从哪个角度来看，我们都需要在本章节中提到的研究领域取得重大进展才能创建更为普遍的人工智能系统[24]。而且，如果我们真的能够取得上述进展，更通用的人工智能方法将给人类带来巨大收益；与此同时，如果安全

挑战未能得到有效应对，它也会带来新的风险。尽管取得突破的几率无法保证，美国仍然应当继续研究更具类人能力的系统，同时进行相应的投资，确保这些系统的安全性和可控性。

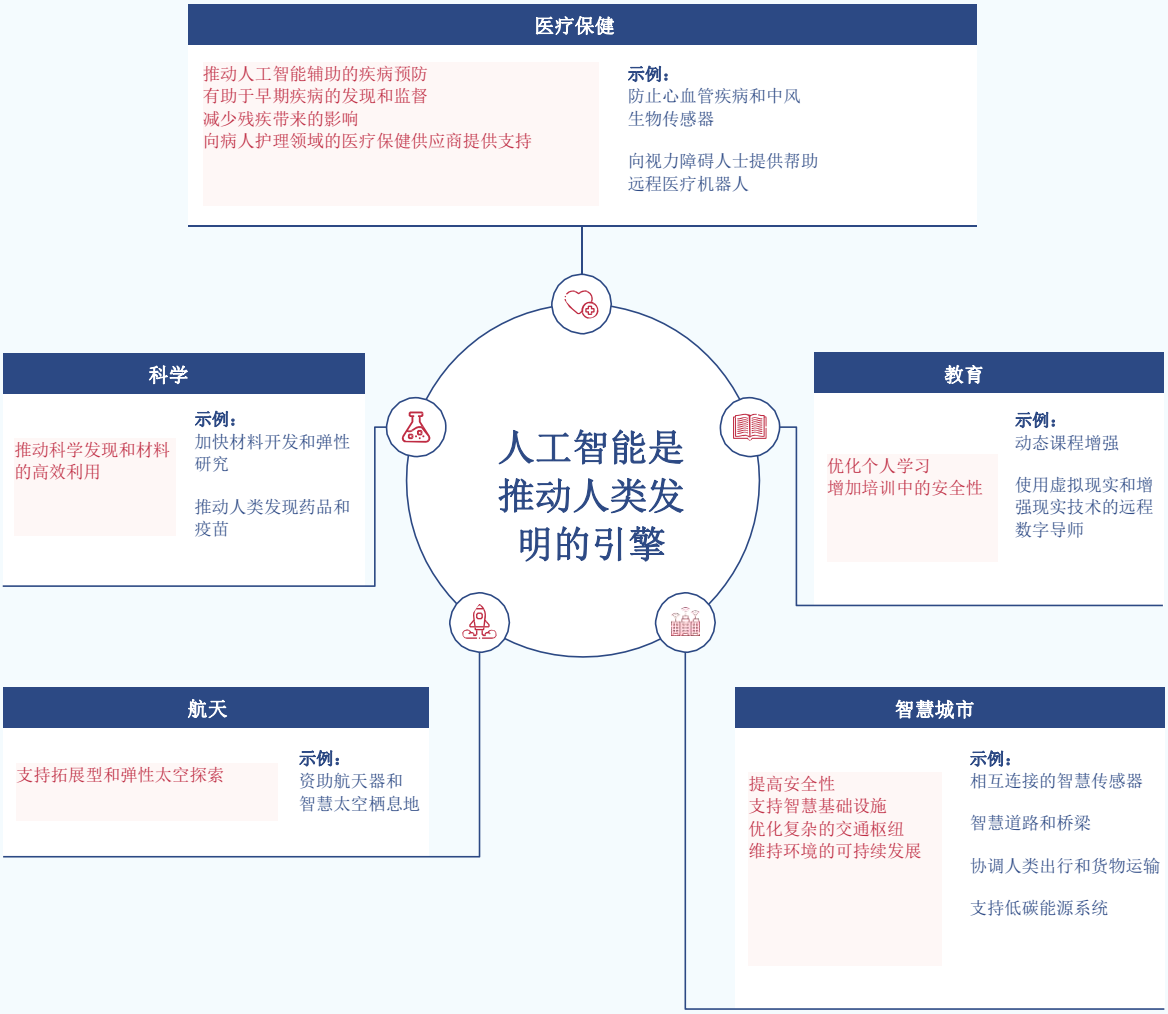
人工智能进步，包括掌握沿着一个或多个维度向前发展的更通用的人工智能能力，将有可能为我们提供新的能力和应用。部分进展可能导致拐点的出现或能力上的飞跃。不过，这些进展也可能引发新的忧虑并带来的新的风险，同时催生对新政策、建议和技术进步的要求，从而确保人工智能系统符合人类的目标和价值观[25]，即符合人类对安全性、稳健性和可信度等的要求[26]。美国应当监控人工智能的发展，开展必要的技术投资，密切关注政策变化，确保人工智能系统符合美国的目标和价值观[27]。

展望人工智能的未来

沿着上述研究线索的轨迹，我们可以勾画人工智能未来的美好愿景：人工智能将以前所未有的方式对人类进行赋能，解锁人工智能在科学、教育、航天技术、医疗保健、基础设施、制造业、农业、娱乐业以及其他行业的能力。举例来说，基于自然语言理解领域内的技术进步，我们可以对书面和口头训练数据极为有限的晦涩语言进行实时的普适翻译[28]。这将改变我们跨越地理和文化障碍进行交流的方式，促进商业、外交和思想的自由交流。

联邦政府在应对流行性疾病和自然灾害时，借助多模态和多源数据融合方面取得的突破性进展，可以于进行实时的、人工智能驱动的建模和仿真活动[29]。通过地图、建筑布局和其他视觉数据层增强的无人机反馈，可以使最先响应者了解急救场景[30]，而人工智能可以协助制定相应计划、加快指挥和控制流程、优化一系列灾难响应场景中的物流服务[31]。

人工智能是推动
人类发明的引擎



语境中的人工智能 - 尾注

[1] 约翰 • 麦凯恩（John S. McCain）2019 财政年度国防授权法案包括以下定义，以指导委员会的工作：1.任何人工系统，在不同的和不可预测的情况下执行任务，而不需要大量的人类监督，或者在接触到数据集时能从经验中学习并提高性能。2.一种在计算机软件、物理硬件或其他环境中开发的人工系统，可解决需要类似人类的感知、认知、计划、学习、交流或身体行动的任务。3.一种人工系统，旨在像人一样思考或行动，包括认知体系结构和神经网络。4.一组技术，包括机器学习，旨在近似于认知任务。5.一种旨在合理行动的人工系统，包括一个智能软件代理或具身的机器人，利用感知、计划、推理、学习、沟通、决策和行动来实现目标。参见出版授权号 115 - 232，132 法令 1636，1965 (2018)。

[2] 安德鲁 • 沃尔（Andrew W. Moore）等人，*人工智能堆栈：*开发和部署人工智能的蓝图，程序 SPIE 10635，持久 ISR IX 的地面/空中多传感器互操作性、集成和联网，106350C（2018 年 5 月 4 日），https://doi.org/10.1117/12.2309483；另见戴夫 • 马丁内兹（Dave Martinez）等人，人工智能：*短暂历史、现在发展，以及未来的展望*，麻省理工学院林肯实验室，27，（2019 年 1 月），https://www.ll.mit.edu/media/9526。

[3] 请注意，基于模型的人工智能需要手动构建模型的数据。通常，这涉及的数据比统计机器学习少，但需要更多的人力。

[4] 萨利玛 • 阿梅尔希（Saleema Amershi）等人，*机器学习软件工程：案例研究*，ICSE-SEIP '19 第 41 届国际软件工程会议论文集，291-300（2019），https://2019.icse-conferences.org/details/icse-2019 -Software-Engineering-in-Practice/30/Software- Engineering-for-Machine-Learning-A- Case-Study；斯库尔利（D. Sculley）等人，*机器学习：技术债务的高利息信用卡*，SE4ML：机器学习软件工程（神经信息处理系统 2014 研讨会），https://ai.google/research/pubs/pub43146.

[5] 约翰 • 劳克伯里（John Launchbury），*美国国防部高级研究项目局对人工智能的看法*，DARPA，4-7（2017 年 2 月），https:// www.darpa.mil/attachments/AIFull.pdf。

[6] 以今天的统计机器学习为例，其局限性包括：容易在不知不觉中学习和放大训练数据中的偏差；通常是由非常多的学习参数组成的复杂模型，使其不透明且难以解释；进行训练，以解决狭义的任务，对其他相关问题缺乏概括性（比如当操作上遇到的数据与训练数据的特征发生根本性变化时）；及其需要大量的标注训练数据。

[7] 安德鲁 • 莫尔（Andrew Moore），*当人工智能成为一种日常技术*，《哈佛商业评论》（2019 年 6 月 7 日），https://hbr.org/2019/06/when-ai-becomes-an-everyday-technology.

[8] 汤姆 • 布朗（Tom B. Brown）等人，*语言模型是少数的学习者*，阿奇夫论文预印本网站（2020 年 7 月 22 日），https://arxiv. org/abs/2005.14165.

[9] 扎卡里 • 阿诺德（Zachary Arnold）等人，*追踪人工智能投资：私营市场的初步发现*，安全和新兴技术中心（2020 年 9 月），https://cset.georgetown.edu/wp-content/uploads/ CSET-Tracking-AI-Investment.pdf.

[10] 据联合国教科文组织称，全球研发支出已达到创纪录的最高水平，接近 1.7 万亿美元。《看看你们国家在研发方面投入了多少？》，联合国教科文组织统计研究所（上次访问时间：2021 年 1 月 7 日），http://uis.unesco.org/apps/visualisations/research-and-development- spending/。

[11] 关于人工智能专家最近的辩论，参见*人工智能辩论 2：推进人工智能：跨学科方法*，蒙特利尔人工智能公司（2020 年 12 月 23 日），https://montrealartificialintelligence.com/ aidebate2.html。

[12] See e.g.，Bryan Wilder，et al.，*学会与人类互补*，人工智能组织国际联合会议（2020 年），https://doi.org/10.24963/ijcai.2020/212；埃塞 • 卡玛尔（Ece Kamar）等人，*在大规模众包中结合人类和机器智能*，第 11 届自主代理和多代理系统国际会议论文集（AAMAS 2012）(2012 年 6 月 4 日至 8 日)，https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/galaxyZoo.pdf；拉米亚 • 拉马克里什南

（Ramya Ramakrishnan）等人，*克服现实世界中的盲点：利用互补的能力进行联合执行*，人工智能促进会的人工智能会议论文集（2019 年 7 月 17 日），https://doi.org/10.1609/aaai.v33i01.33016137；萨利玛 • 阿梅尔希（Saleema Amershi）等人*人类与人工智能互动的准则*，CHI ’ 19：2019 年 CHI 计算系统中的人的因素会议论文集（2019 年 5 月)https://doi.org/10.1145/3290605.3300233；埃里克 • 霍维茨（ Eric Horvitz），关于混合发起式互动的挑战和承诺的反思，人工智能杂志（2007 年 6 月 15 日）https://doi. org/10.1609/aimag.v28i2.2036。

[13] 这些新方法的一个目标是消除对许多复杂计算的需要，这些计算使传统的训练非常缓慢。文森特 • 迪托杜瓦尔（Vincent Dutordoir），*具有球面谐波特征的稀疏高斯过程*，阿奇夫论文预印本网站（2020 年 6 月 30 日），https://arxiv.org/abs/2006.16649.

[14] 关于结合了符号操作和深度学习的混合智能架构的讨论，请参阅加里 • 马库斯（Gary Marcus）的人工智能的下一个十年：*Four Steps Towards Robust Artificial Intelligence*，阿奇夫论文预印本网站 at 14-19（Feb. 19，2020），https://arxiv.org/abs/2002.06177.

[15] 见*神经象征性人工智能*，麻省理工学院-IBM 沃森人工智能实验室（上次访问时间：2021 年 1 月 16 日），https:// mitibmwatsonailab.mit.edu/category/neuro-symbolic-ai/。

[16] 阿努杰 • 卡帕特内（Anuj Karpatne）等人，*物理学指导下的神经网络（PGNN）：湖泊温度模拟中的应用*，计算机协会知识发现和数据挖掘特别兴趣小组（）2018（2018 年 2 月 20 日），https://arxiv.org/pdf/1710.11431.pdf。

[17] 雷纳 • 拉贾特（Rajat Raina）等人，*自学成才：《从无标签数据中转移学习》*，第 24 届国际机器学习会议的会议记录（2007 年 6 月），https://dl.acm.org/doi/ ABS/10.1145/1273496.1273592；布鲁斯 • 德雷珀（Bruce Draper）博士，*学习较少标签*，美国国防部高级研究项目局（上次访问时间：2020 年 12 月 19 日）https://www.darpa.mil/program/learning-with-less-labeling。

[18] 拉胡尔 • 迪克西特（Rahul Dixit）等人，稀疏/不精确数据环境下的人工智能与机器学习，电气与电子工程师协会（2020 年 8 月 21 日），https://ieeexplore.ieee.org/document/9172612.

[19] 参见杰姆 • 迪米加尼（Cem Dilmegani）*2021 年合成数据终极指南*，AI Multiple（Jan. 12，2021），https://research.aimultiple.com/synthetic-data/.

[20] *合成数据的真正潜力*, 麻省理工学院新闻（2020 年 10 月 16 日），https://news.mit.edu/2020/real- promise - synthetic - data -1016.

[21] *10 项突破性技术* 2020 年，麻省理工学院技术评论（2020 年 2 月 26 日），https://www. technologyreview.com/10 -breakthrough-technologies/2020/#tiny-ai。

[22] *全球最大最全的常识性知识库*，Cycorp 公司（上次访问时间：2020 年 12 月 19 日），https://www.cyc.com/the-cyc-platform/the-knowledgent-base；约翰 • 帕夫卢斯（John Pavlus），*常识更接近电脑*，广达杂志（2020 年 4 月 30 日），https：//www.quantamagazine。org/common-sense-comes-to-computers-20200430/；安托万 • 博塞鲁特（Antoine Bosselut）等人，*COMET：自动知识图构建的常识性变形，第 57 届计算语言学协会年会论文集（2019）*https://homes.cs.washington.edu/~msap/pdfs/bosselut2019comet.pdf。

[23]艾米 • 布卢门撒尔（Amy Blumenthal），如何让人工智能可信，《每日科学》（2020 年 8 月 31 日），https://www. sciencedaily.com/releases/2020/08/200827105937.htm。

[24]参见本尼迪克特 • 尼奥（Benedict Neo），领先于人工通用智能竞赛的四大人工智能公司，《走向数据科学》（2020 年 4 月 13 日）,https://towardsdatascience.com/four-ai-companies-on- the-bleeding-edge-of-artificial-general-intelligence-b17227a0b64a；薛瑞希蒂 • 迪奥拉斯（Srishti Deoras）, *9 家公司在 AGI 领域做着卓越的工作，就像 OpenAI 一样*，印度分析杂志 (July 25, 2019), https:// analyticsindiamag.com/9 -companies-doing-exceptional-work-in-agi-just-like-openai/。

[25]在“关键考虑因素”中，委员会描述了开发和实施符合关键价值的系统的做法、技术和操作政策。重要的是，各机构必须考虑价值，(1)体现在关于工程权衡的选择中，(2)明确体现在人工智能系统的目标和效用函数中。请参阅*《负责任开发和人工智能的核心考虑*因素》，人工智能国家安全委员会 (2020 年 7 月)，https://www.nscai.gov/wp-content/uploads/2021/01/ Key-Considerations-for-Responsible-Development-Fielding-of-AI.pdf。除了现在需要的行动和投资（见本报告第 7 章相关的行动蓝图），关键考虑因素包括随着技术的发展应该更新的政策和做法，以反映新的人工智能考虑。

[26]如本报告的关键考虑因素和第七章所述内容，则将需要研发。它还需要继续投资于系统架构，以便限制系统故障的后果，在系统运行时监测人工智能的性能，用以评估其是否按预期执行，并克服审计和监督方面的科技差距。有关详细信息，请参阅本报告第 7 章和第 8 章。

[27]如第 8 章及其行动蓝图所述，例如，这将需要对人工智能应用和新兴技术的隐私和公民自由的影响进行持续评估和经常性指导。随着社区规范和技术能力的变化，将需要不断更新不可接受的结果和政策指导。此外，“关键考虑因素”指出，工程实践将需要评估一般的可行性和遵守政策中所表达的不允许的结果，以及具体的候选人工智能技术的技术成熟度。请参阅《负责任开发和人工智能的核心考虑因素》，人工智能国家安全委员会 (2020 年 7 月)，https://www.nscai.gov/wp-content/uploads/2021/01/Key- Considerations-for-Responsible- Development- Fielding - of- AI.pdf.

[28]设想中的翻译系统将通过积极纠正翻译和软件识别错误来利用对系统的反馈，并随着翻译各方之间的互动不断进行提高性能。一个主要目标是快速部署到以前从未见过的新语言。例如，卡内基梅隆的外交官项目通过名为多引擎机器翻译（MEMT）的新架构实现了交互式语音翻译。外交官项目使用户能够提供翻译更正功能，以便支持快速开发以前未从未见过的新语言。罗伯特 • 弗雷德金，*外交官项目中的交互式演讲翻译*，卡内基梅隆大学语言技术学院（最上次访问时间：2020 年 12 月 19 日），http://www.cs.cmu.edu/~air/papers/acl97-workshop.pdf。

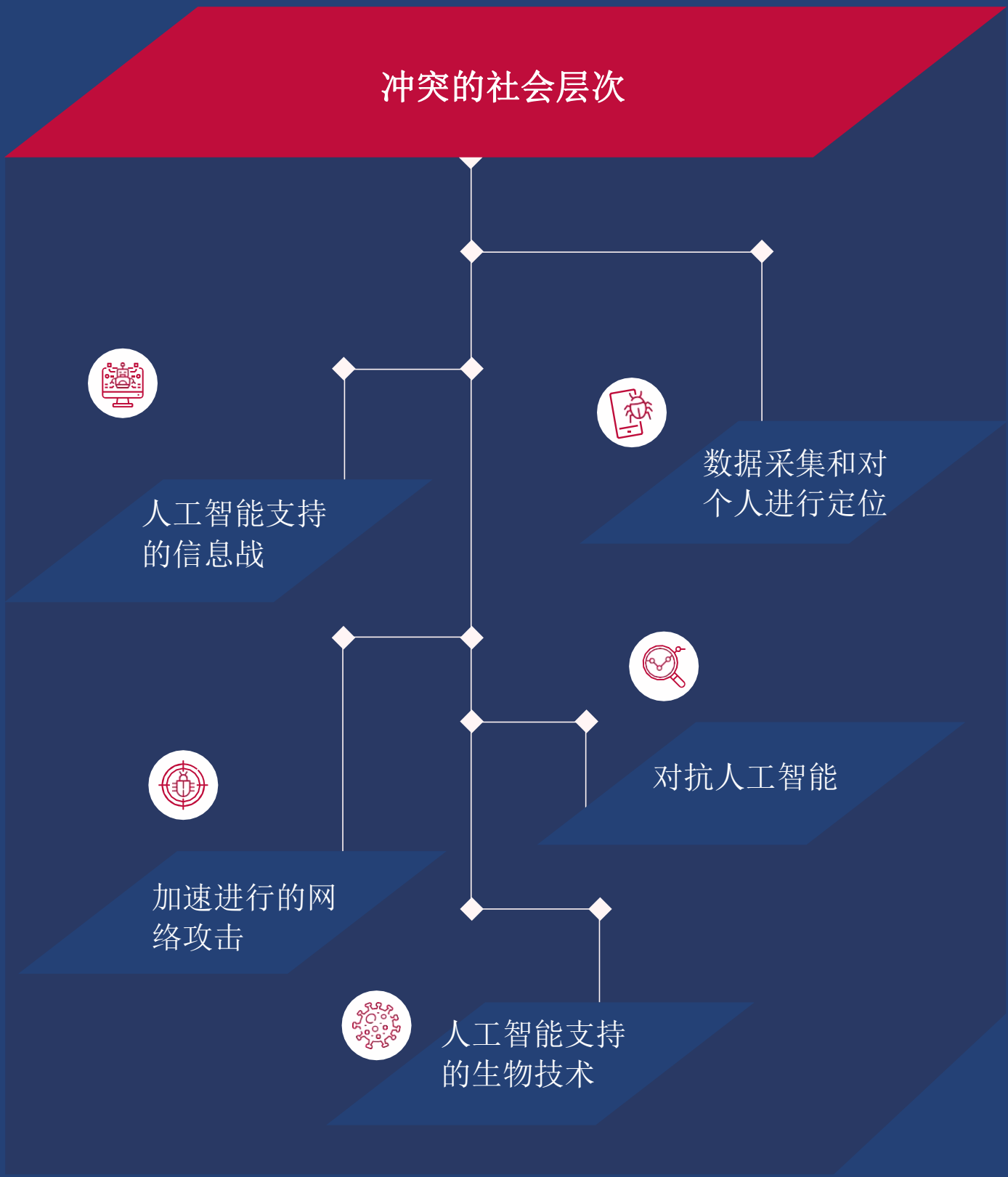
[29]建模和模拟还可以帮助为政府精心策划的有效的大流行病供应链对策做好准备。马达夫-马拉特（Madhav Marathe），*《支持实时 COVID19 响应的高性能模拟》*，SIGSIM-PADS ’ 20（2020 年 6 月），https://dl.acm.org/doi/ pdf/10.1145/3384441.3395993。

[30] Edgybees 公司（上次访问时间 2020 年 12 月 19 日），https://edgybees.com/。

[31]*美国能源部宣布成立第一个五大联盟*，美国能源部（2020 年 8 月 18 日），https://www.energy.gov/articles/department-energy-announces-first-five-consortium

第一部分

第一章 人工智能时代的新兴威胁


冲突的社会层次
人工智能支持的信息战
数据采集和对个人进行定位
对抗人工智能
加速进行的网络攻击
人工智能支持的生物技术

美国政府对于如何在即将到来的人工智能时代保卫国家安全尚未做好准备。目前，人工智能应用正在对现存威胁进行改头换面，创造新的威胁类别，进一步鼓励国家和非国家对手利用我们开放社会中存在的漏洞[1]。正如导弹时代和恐怖主义使威胁距离美国本土只有一步之遥，人工智能系统将使对手的射程和影响力深入美国本土。借助人工智能技术，对手能够以微观精度、宏观尺度和更快的速度采取行动。它们将利用人工智能加强网络攻击和数字转型，并用新方法对个人进行定位。人工智能还有助于制造精确设计的生物制剂。此外，对手将操纵我们依赖的人工智能系统。

人工智能技术加剧了现存的两大国家安全挑战：

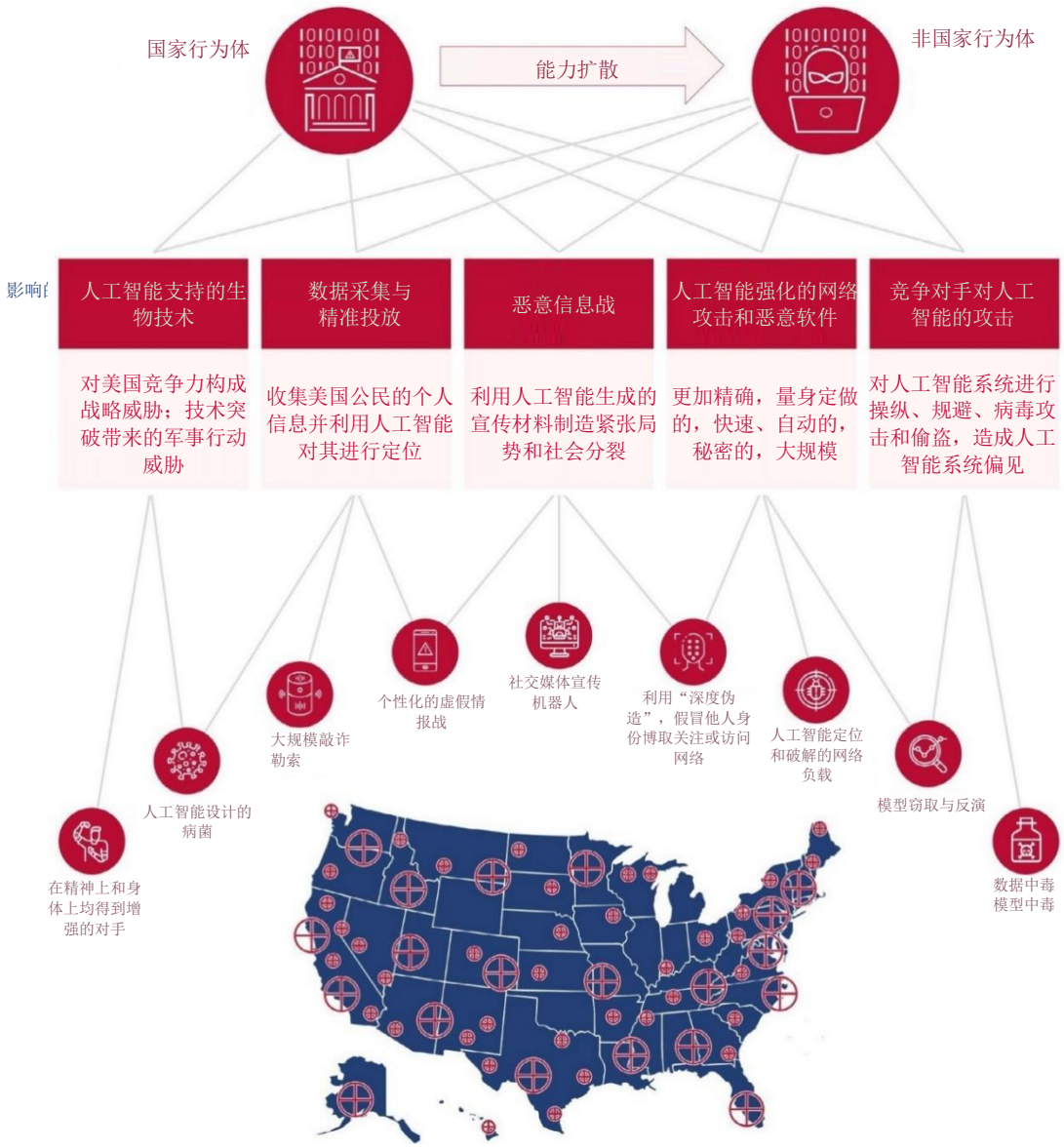
- 第一，在现实生活中，各行各业形成的数字依赖使得公司、大学、政府、私人组织和家庭等社会组成部分都更容易受到网络入侵。与此同时，新型传感器已经遍布现代世界的每个角落。物联网、汽车、电话、房屋和社交媒体平台等持续收集数据流，并将其输入人工智能系统，对手将利用这些数据流对我们的公民进行识别、定位、操纵或胁迫[2]。

- 第二，国家和非国家对手正在利用网络攻击、间谍活动、心理战和政治战、金融工具等不会造成直接军事对抗的手段对美国进行挑衅。它们无需人工智能就可以

发动大规模网络攻击、窃取关于美国公民的海量敏感数据、干涉我们的选举，或在数字平台上利用恶意信息对我们进行轰炸。但是，人工智能正在改变这些攻击的种类和程度，给美国的经济、关键基础设施和社会凝聚力带来新的威胁[3]。这些人工智能能力将用于冲突的整个范围。而且，它们将成为非军事对抗首选的工具，军事行动的序幕，或与战争中的军事行动同时进行。

美国民众已经开始意识到这种数字依赖的隐私影响及“深度伪造”等人工智能支持的恶意信息带来的潜在威胁。尽管如此，在美国境内围绕人工智能进行的这场辩论仍然未能涵盖人工智能威胁的全部范围和威胁，以及我们周围的人工智能系统面临的总体安全风险。对手将通过机器学习、规划和优化等手段创建系统，进而以无法察觉的方式对美国公民的信仰和行为进行操纵。鉴于这种可能性，我们不得不承认美国正处在暴风雨的前夕[4]。更让人担忧的是，对手利用人工智能建造大规模杀伤性武器；在未来的战争中，每一个公民和组织都可能成为这些武器潜在的打击对象。.

“对手将通过机器学习，规划和优化等手段创建系统，进而以无法察觉的方式对美国公民的信仰和行为进行操纵。鉴于这种可能性，我们不得不承认美国正处在暴风雨的前夕 。更让人担忧的是，对手利用人工智能建造大规模杀伤性武器，在未来的战争中，每一个公民和组织都可能成为这些武器潜在的打击对象。”



本章余下的部分将探讨对手已经完成开发并已经被用来攻击美国，或即将开发并将被用来攻击美国的五大人工智能威胁。

1. 工智能支持的信息战

人工智能及相关技术将增加对手信息战的影响、精准度和持续性。人工智能从以下三个方面加剧了恶意信息问题：

- **信息。**人工智能可以产生原始的基于文本的内容，并通过基于生成对抗网络和强化学习的“深度伪造”技术等手段，对图像、音频和视频进行处理，从而造成很难与真正信息进行区分。

- **受众。**人工智能可以对个人的偏好、行为和信仰进行侧写，以便定位特定的群体受众，并向其发送有针对性的信息。

- **媒介。**人工智能可以嵌入在平台内部，并且通过排名算法等手段，传播恶意信息。

与 20 世纪的宣传活动不同，人工智能支持的恶意信息战不仅仅是向 100 万人发送一条强有力的信息，而是在对目标群体的数字生活、情绪状态和社交网络进行详细解读的基础上，分别发送 100 万条个性化定制的信息[5]。目前，竞争国家已经开始利用人工智能支持的恶意信息，包括依赖人工智能生成的虚假身份信息[6]。控制和操纵数字信息对于俄罗斯政府的战略至关重要，目前它已经在美国以及世界其他国家和地区开展了一系列试图破坏民主进程完整性的活动[7]。

在美国，私人部门在对抗外国恶意信息方面发挥了重要的领导作用。特别值得一提的是，社交媒体公司在旗下的平台上进行了声势浩大的信息追踪和管理工作。但是，政府与社交媒体公司之间的协调关系仍然是临时性的。因此，针对外国生成的虚假信息问题，我们需要一个全面统筹的公共-私人应对机制。而且，政府需要投入更多的精力和资源来应对检测、归属和媒体认证带来的技术挑战。政府应当：

建议

*组建一个联合跨部门的特别工作组和行动中心。*目前，国会已经授权在国家情报总监办公室成立外国恶意影响应对中心[8]。此外，政府还应通过新成立的权力机构，组建技术上先进的、24 小时不间断作业的工作组和行动中心，领导和整合政府部门针对源自海外的恶意信息采取的反击行动。中心将对相关的公共和私人部门的工作场景进行考察，在双方之间开展协调工作，并且及时针对外国发动的信息战及时进行反击。为了高效进行曝光、归因及应对工作，中心必须配备现代化的人工智能数字工具及具备专业知识的各类人才。

建议

*向 DARPA 提供资金，使其能够在多个旨在检测、归因和破坏人工智能支持的恶意信息活动及验证数字媒体来源的研究项目之间进行协调。*额外的资金支持不但能够扩大 DARPA 当前进行的旨在检测合成媒体研究项目的规模，而且还使这些项目的研究范围扩展至对恶意信息活动的归因及破坏[9]。尽管部分检测技术前景光明，但是资助开发验证数字媒体来源的替代性技术，将为我们提供技术上更为稳健的手段来防止假冒可信的信息来源[10]。DARPA 应努力完成这些项目，协助政府部门和机构进行相关技术转化和应用，确保美国政府能够对恶意信息活动进行实时检测、归因和破坏。

建议

*组建特别工作组来研究人工智能和补充技术的使用，包括标准和技术的开发与部署，以便对内容的真实性和来源进行验证。*白宫科技政策办公室应牵头组建这个特别工作组。为了应对错误信息带来的挑战，相关部门目前正在积极开发旨在验证视听内容真实性和来源的标准和渠道[11]。开发工作利用包括加密和脆弱水印在内的各项技术，通过内容生产和传输渠道，保护和跟踪内容的预期转换。与此同时，开发工作为减少恶意信息活动创造了条件，这些恶意活动寻求对我们的数字生态系统中受到高度信任的信息源进行破坏和欺骗。目前，在该技术领域内建立公共-私人部门间的合作伙伴关系的时机已经成熟。几家私人组织已经着手开始组建，打击该领域内的虚假信息活动[12]。

2. 数据采集和对个人进行定位

“潜在对手将充分认识到广告公司和社交媒体企业‘制胜法宝’的威力：人工智能是一款非常强大的定位工具。”

数据安全也关系着国家安全。今天，“广告技术”已经成为“国家安全技术”。潜在对手将充分认识到广告公司和社交媒体企业“制胜法宝”的威力：人工智能是一款非常强大的定位工具。过去，人工智能支持的分析学改变了公司与客户之间的关系；今天，它改变了政府与个人之间的关系。个人数据的广泛流通在推动商业创新的同时，也带来了很大的漏洞[13]。我们担心，对手正在有计划有组织地收集关于美国企业、公民和政府的数据，而这已经大大超出了传统间谍活动的范畴[14]。对手将广泛可用的商业数据与非法获取的数据（例如，2015 年，黑客对美国人事管理办公室发动了攻击，获取了大量雇员信息）结合起来，从而对个人进行跟踪、操纵和胁迫[15]。在缺乏充分的数据保护时，人工智能使得个人很难隐藏自身的财务状况、生活方式、人际关系、健康状态，甚至情绪变化。随着对手开始描绘个人、网络和社会分裂状态，预测对不同刺激因素的反应及模拟最佳操纵行为或伤害行为，个人和商业漏洞已经成为国家安全薄弱环节。与此同时，这些技术的兴起与传播也带来了严峻的反情报挑战[16]。

为了推动政府将公民和企业数据上升到国家安全资产的战略高度，我们需要在数据安全思维方式及政策和法律层面进行重大调整来加强数据安全管理；我们需要明确最敏感的个人和商业数据的类别与组合；我们需要从开始阶段就限制外国对手进行数据收集工作，例如联邦政府担心敏感的私人数据被恶意使用，要求一家中国企业放弃一款流行的约会应用程序的所有权[17]，这代表政府从源头上遏制数据采集的一项重大举措。但是，政府缺少整合了明确的政策、清晰的标准和主管部门职责划分的综合性方法，来应对这个涉及多方面的问题。政府应当：

制定和出台将数据安全作为国家安全重要环节的相关政策，其中应包含下列领域：

- **第一，从技术角度来看，政府必须确保人工智能系统（包括收购的商业系统）全生命周期安全开发方法已经到位**，其中重点关注潜在的隐私攻击[18]。“红队测试”必须包括针对隐私进行的专家鉴定流程。政府数据库必须尽可能地进行联合和匿名化处理，个人数据保留周期不得超过必要时限，从而增加对手将信息用于恶意用途的难度。

- 第二，政府应当确保数据隐私和安全性是加强外国投资筛选、供应链情报和风险管理大型计划中优先考量的因素[19]。

- 第三，在国家层面对数据保护和隐私进行立法和监管时，必须结合国家安全考量因素，例如限制不友好的外国行为体在商业市场上获取关于美国民众敏感信息的能力[20]。

3. 加快针对网络攻击进行布局

在人工智能时代，一旦恶意软件植入计算机系统之后，它可以变异成数千种不同的形式。目前，这种变异的多态性恶意软件已经占恶意可执行文件的 90%以上[21]。深度强化学习工具已经能够发现漏洞，隐藏恶意软件，并且有选择性地进行攻击[22]。虽然哪种方法占据主导地位仍不明确，但是摆在美国对手面前的却是一条清晰的道路：它们可以利用一系列新旧算法手段组合来实现攻击的自动化、优化和自动通知，从而改变网络攻击和间谍活动的有效性[23]，而这已经超越了人工智能增强型恶意软件的范畴。在网络攻击活动的所有阶段，机器学习均有现前和潜在的应用价值，并且它还将改变网络战争和网络犯罪的本质[24]。随着现有人工智能网络能力的应用范围不断扩大，网络攻击将呈现出更加精确和更具个性化的特质，进一步加速网络战争，使网络战争开始朝着自动化的方向发展，使网络武器变得更具隐秘性和持久性，使网络战的规模更大，攻击效果更佳。

“在网络攻击活动的所有阶段，机器学习均有现前和潜在的应用价值……”

事实证明，美国的防御体系甚至无法对最基本的网络挑衅做出有力的回应。时至今日，我们过时老旧的基础设施和医疗设备仍然存在着公开的漏洞，而 5G 网络、数十亿物联网设备和软件供应链中的新漏洞正在不断激增[25]。2017 年，俄罗斯发动了 NotPetya 网络攻击，造成全球损失高达数十亿美元，这充分证明了基础级别的自动化恶意软件的威力，有能力的国家行为体的风险承受能力以及这种能力扩散带来的后果[26]。虽然人工智能的防御性应用为提高国家网络防御实力带来了希望，但是由于数字基础设施的天生脆弱性，人工智能无法向其提供有效保护。为了解决当前威胁，国会必须继续执行网络空间日光浴室委员会提出的建议[27]。在网络防御基础夯实之后，美国可以通过测试和构建人工智能支持的网络防御系统所需的仪器化基础设施、建立更加切实可行的安全奖励措施、妥善组织迎接挑战以及持续使攻击者处于失衡状态，为不断扩大的威胁做好准备。与此同时，除非联邦政府采取紧急行动，否则渗透性网络间谍活动及对美国计算机网络和关键基础设施的网络攻击仍将继续，并且将更具破坏性。因此，政府应当：

建议

开发和部署人工智能防御系统来抵御网络攻击。国家安全机构需要采购必要的传感器和仪器，以便训练人工智能系统检测和应对自身网络受到的威胁。与此同时，需要对人工智能防御系统进行大规模仪器化的真实测试。防御系统必须足够稳健，能够有效抵御对抗攻击。国家安全机构应当在政府网络内大规模部署防御系统，扩大以机器速度运行的信息共享和基于行为的异常检测，同时减少恶意软件带来的干扰和破坏。为了充分利用上述能力，政府应当加快以国家反恐中心为蓝本的联合网络规划和运营中心的组建工作[28]。该中心将作为集中式网络情报共享和协作单位，由多个机构进行管辖，由主管部门调查威胁，积极支持防御缓解措施，并协调应对措施。

4. 对抗人工智能

人工智能系统代表着全新的攻击目标。尽管我们正处在这一新现象的前沿，商业公司和研究人员已经记录了涉及规避、数据中毒、模型复制和利用传统软件缺陷进行欺骗、操纵、破坏和使人工智能系统失效的各种攻击[29]。上述威胁与传统的网络活动相关，不过二者之间存在明显的差异。当人工智能系统加强某个域内的平民或军事行动时，它很容易受到来自该域的对抗攻击[30]。鉴于人工智能系统对大数据集和算法的依赖，即便对这些数据集或算法进行微小的操作，也将导致人工智能系统的操作出现重大变化。这种威胁并不只是一种假设：对抗攻击正在发生并且已经影响了商业性机器学习系统的运行[31]。除了少数例外情形，由于研发投资不足，保护人工智能系统一直是设计和部署人工智能系统过程中的“事后想法”[32]。在最近对 28 家组织进行的一项调研中，仅有三家受调研组织拥有“保护机器学习系统的正确工具”[33]。目前，政府尚未采取统一的行动在整个国家安全企业内部整合人工智能保障工作。为了提高人工智能“保障”，政府应当：

建议

建立国家人工智能保障框架。所有政府机构都需要开发和应用对抗性机器学习威胁框架，解决关键人工智能系统遭遇到的攻击和应采取的防御行动问题。分析框架有助于对政府人工智能系统面临的威胁进行分类，并且能够协助分析人员对威胁和漏洞进行检测、应对和采取相应的补救措施[34]。

建议

组建专门进行对抗性测试的红队。红队应当采取进攻姿态，努力尝试破坏系统并且使系统违反正确的行为规则。鉴于人工智能红队队员所需的专业知识和经验的稀缺性，国防部和国家情报总监办公室应当考虑建立跨政府部门的人工智能红队能力社区，并在社区内进行多种人工智能开发[35]。

5. 利用人工智能，支持生物技术发展

今天，生物学已经成为一门可以进行编程操作的学科。基因编辑工具 CRISPR 等新技术的出现，拉开了人类对 DNA 进行编辑的序幕。基于强大的计算能力和人工智能，生物技术创新或可为长期困扰人类的挑战提供全新的解决方案，例如健康、食品生产和环境可持续发展等问题。不过，与其他强大的技术一样，生物技术的应用也有黑暗的一面。新冠肺炎肆虐给全世界敲响了警钟，提醒人类留意高度传染性病原体的危险。借助人工智能，人类可以设计具有致命杀伤力的病原体，或使病原体感染具备某一基因档案的特定人群，从而使生物技术变成“终极杀招”。此外，当人工智能应用于生物领域时，它可以从智慧上和身体上对人类进行优化，实现生理机能的强化。在某种程度上，由于脑电波可以表现为对人工智能的机器视觉挑战，因此大脑的秘密可能就此揭开——人类甚至可以对大脑进行编程。

对此，个人、团体和国家将会持有不同的道德和伦理观点，并且以进步的名义，愿意承担不同程度的风险。相对而言，美国的竞争对手很可能采取风险容忍行动，而不是严格遵守生物技术规范和标准。中国深知引领生物革命浪潮能够带来巨大的优势。华大集团（前身为北京基因组研究所）等企业和科研机构的海量基因组数据集，以及中国目前的全球性基因数据收集平台和“属于全人类”的人工智能研究方法，将会使中国成为生物技术领域内令人敬畏的竞争对手[36]。坦白地说，美国无法承受在 10 年之后，当它回顾生物技术发展进程时，“惊讶地”发现在生物技术领域出现一家地位和实力相当于华为的中国企业的代价。此外，俄罗斯长期无视科学规范和生物伦理原则。具体来说，它研制和使用 Novichok 等新型神经毒剂用于暗杀目的，我们对于俄罗斯政府遵守《禁止生物武器公约》的情况表示忧虑。与此同时，这也预示着俄罗斯政府有意愿将先进的生物技术能力用于极其恶毒的目的[37]。为此，政府应当：

建议

增加美国国家安全机构对生物安全和生物技术问题的关注度。鉴于人工智能将大大提升生物技术进步的速度，政府应当更新《国家生物防御战略》，以便将更大范围内的生物威胁（例如人类机能强化）、基因数据的恶意用途以及出于对其他新用途的考虑，竞争对手利用生物技术或生物数据优势的方式纳入进来。此外，美国官员应当就外国行为体获取个人基因信息带来的威胁发出警告，重点强调华大集团和中国政府引发的担忧[38]。

第一章 – 尾注

[1] 威胁可以理解为敌方的能力与可能造成有害后果的弱点相搭配。请参阅特里 • 德贝尔（Terry L. Deibel）的《*外交策略：美国治国之道的逻辑*》，剑桥大学出版社，142-150（2007）。可根据严重性、可能性、迫切性和可处理性对威胁进一步分级。

[2] 斯图尔特 • 汤普森（Stuart A. Thompson）和查理 • 瓦策尔（Charlie Warzel），*1200 万部电话，一个数据集，零隐私*，《纽约时报》（2019 年 12 月 19 日），https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking- cell - phone.html。例如，物联网（IoT）和人工智能驱动的应用可以将你的新扫地机器人变成一个监听器。请参阅斯里拉姆 萨米（Sriram Sami）等人，扫地机器人的*间谍活动：通过激光雷达传感器进行窃听*，第 18 届嵌入式网络传感器系统会议论文集（2020 年 11 月），https://dl.acm.org/doi/10.1145/3384419.3430781。

[3] 这在某种程度上类似于冷战时期战略家所说的“反价值定位”。参阅劳伦斯 • 弗雷德曼的《核战略的演变》，麦克米兰出版公司（Palgrave Macmillan）第 20 卷，119 -122 (1989)。在核战略领域，这也之称为反城市或反经济的目标。

[4] 一些观察者用“锐实力”的概念来描述这种在开放社会中施加影响的努力。这些权力的使用是尖锐的，“因为[专制国家目的是]刺穿、渗透或穿透目标国家的信息环境”。*锐实力：专制主义的影响力不断上升，国家民主基金会*，13，（2017 年 12 月 5 日）https://www.ned.org/sharp-power-rising-authoritarian-influence-forum-report/。另见微软公司埃里克 • 霍维茨（Eric Horvitz）博士在美国参议院商业、科学和运输委员会空间、科学和竞争力小组委员会的证词，关于人工智能黎明的听证会，13，（2016 年 11 月 30 日），http://erichorvitz.com/Senate_Testimony_ Eric_ Horvitz.pdf.

[5] 一些人将人工智能驱动的信息操作描述为“计算宣传”。参见马特 • 切森的 *MADCOM 未来：人工智能将如何加强计算宣传、重新编程人类文化和威胁民主……以及如何应对这一问题*，大西洋理事会 (2017 年 9 月), https://www.atlanticcouncil.org/wp-content/uploads/2017/09/ The_ MADCOM_ Future_ RW_0926.pdf.

[6] 詹姆斯 • 文森特（James Vincent），一场网络宣传活动使用人工智能生成的头像来制造假记者，The Verge 网（2020 年 7 月 7 日），https://www.theverge.com/2020/7/7/21315861/ai-generated- headshots-profile-pictures-fake-journalists-daily-beast-investigation.

[7] 关于俄罗斯干涉 2016 年大选的技术方面的最新研究，请参阅亚历山大 • 斯潘格（Alexander Spangher）等人，*《描述搜索引擎对互联网研究机构网络资产的流量》，Web 会议 (2020 年)*，https://www.microsoft.com/en-us/research/publication/characterizing- search-engine-traffic-to-internet-research-agency-web-properties/；Ryan Boyd 等人，*《使用语言分析描述互联网研究机构在 2016 年美国总统选举期间的社交媒体运营》*，PsyArXiv Preprints (2018)，https://psyarxiv.com/ajh2q/.另见阿琳娜 波利亚科娃（Alina Polyakova），*《弱者的武器：*俄罗斯和人工智能驱动的不对称战争，布鲁金斯学会（2018 年 11 月），https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven- asymmetric-warfare/。

[8] 参见出版授权号 116-92，《2020 财政年度国防授权法》，133 法令 1198, 2129 (2019)。

[9] These include the Media Forensics (MediFor) and Semantic Forensics (SemaFor) programs. See Dr. Matt Turek, *Media Forensics*, DARPA (last accessed Jan.10, 2021), https://www.darpa.mil/program/ media - forensics; Dr. Matt Turek, *Semantic Forensics*, DARPA (last accessed Jan.10, 2021), https:// www.darpa.mil/program/semantic-forensics.

[10] 请参阅，保罗 • 英格兰（Paul England）等人的*《AMP：通过证据认证媒体》，阿奇夫论文预印本网站*（2020 年 6 月 20 日），https://arxiv.org/abs/2001.07886。

[11] 请参阅，保罗 • 英格兰（Paul England）等人的《*AMP：通过证据认证媒体*》，阿奇夫论文预印本网站。

[12] 参见创建数字内容归属标准，内容真实性倡议，https://contentauthenticity.org；和*项目来源：保护受信任的媒体*，项目起源，https://www. originproject.info/about。

[13] 罗伯特 • 威廉姆斯（Robert Williams ）介绍了政策制定者如何面对“创新-安全难题”，其中一个方面是“担心数据隐私和国家安全日益相互关联。数据（和数据网络）可以以威胁安全的方式被利用，但它们也构成了技术创新的命脉，而经济增长和国家安全都依赖于此。”罗伯特 • D.威廉姆斯（Robert D. Williams），*拟定多边技术和网络安全政策*，布鲁金斯学会，1，（2020 年 11 月），https://www.brookings.edu/wp-content/uploads/2020/11/Robert-D-Williams.pdf。

[14] 中岛（Ellen Nakashima）*与一系列主要黑客合作，建立了美国数据库*，《华盛顿邮报》（2015 年 6 月 5 日）https://www.washingtonpost.com/world/national-security/in-a-series-of-hacks- china-appears-to-building-a-database-on-americans/2015/06/05/d2af51fa- 0ba3-11e5 -95fd-d580f1c5d44e_story.html。

[15] 另一个敌对方获取美国个人重要数据的例子是信用报告机构 Equifax 遭黑客入侵。司法部新闻稿，*负责计算机欺诈、经济间谍和黑客入侵信用报告机构的电汇欺诈的中国军方人员，Equifax* (2 月)https://www.justice.gov/opa/pr/chinese-military-personnel-charged- computer-fraud-economic-espionage-and-wire-fraud-hacking；阿鲁纳 • 维斯瓦那塔（Aruna Viswanatha）等人，中国军方四名成员因大规模的 Equifax 漏洞被起诉，*《华尔街日报》（2 月 14 日）报道* https://www.wsj.com/articles/four-members-of-china-s-military-indicted-for-massive-equifax- breach-11581346824。

[16] 参见詹姆斯 • 贝克，*人工智能的反智能影响——第三部分*，“法律战计划”组织（2018 年 10 月 10 日），https://www.lawfareblog.com/counterintelligence-implications-artificial-intelligence-part-iii.

[17] 袁阳（Yuan Yang）和詹姆斯 • 丰塔内拉-汗（James Fontanella-Khan），因美国国家安全问题，Grindr 被中国业主出售，金融时报（2020 年 3 月 7 日），https://www.ft.com/content/a32a740a-5fb3 -11ea-8033-fa40a0d65a98。

[18] 关于隐私攻击，参见玛丽亚 • 里加基（Maria Rigaki）和塞巴斯蒂安 • 加西亚（Sebastian Garcia）的*机器学习中的隐私攻击调查*，阿奇夫论文预印本网站（July 15，2020），https://arxiv.org/abs/2007.07646.

[19] 美国外国投资委员会（CFIUS）有权审查包括“可能以威胁国家安全的方式利用美国公民的敏感个人数据”的交易。有关背景，请参阅劳拉 • 杰尔（Laura Jehl），*随着外国投资规则的生效，敏感的个人数据成为关注的焦点*，《国家法律评论》（2020 年 2 月 18 日），https://www.natlawreview.com/article/spotlight-sensitive-personal-data-foreign-investment-rules-take- force。国家反情报与安全中心（NCSC）在其关键供应链风险的概念中包括“敏感的政府数据和个人身份信息”。请参阅《*供应链风险管理：减少对美国主要供应链的威胁*》，国家反情报与安全中心，3 (2020)，https:// www.dni.gov/files/NCSC/documents/supplychain/20200925 -NCSC-Supply- Chain-Risk-Management- tri -fold.pdf。

[20] 例如，参见格雷厄姆 • 艾利森（Graham Allison），*应用程序禁令不会使美国的安全风险消失*，麻省理工学院技术评论（2020 年 9 月 21 日），https://www.technologyreview.com/2020/09/21/1008620/wechat-tiktok-ban-china-us-security-policy-opinion/。

[21] 尼古拉斯 • 杜兰（Nicholas Duran）等人，*《2018 年 Webroot 威胁报告》*，Webroot，6（2018 年），https://www-cdn.webroot. com/9315/2354/6488/2018-Webroot-Threat-Report_US-ONLINE.pdf。

[22] 加里 • 萨维德拉（Gary J. Saavedra）等人，机器学习在模糊检验中的应用回顾，阿奇夫论文预印本网站（2019 年 10 月 9 日），https://arxiv.org/pdf/1906.11133.pdf；高祖勇夫（Isao Takaesu），机器学习安全：DeepExploit，GitHub（2019 年 8 月 29 日），https://github.com/13o-bbr-

bbq/machine_learning_security/tree/master/DeepExploit；马克 Ph.施特克林（Marc Ph. Stoecklin）等人，*DeepLocker：人工智能如何为隐蔽的新品种恶意软件提供动力*，安全情报网站（2018 年 8 月 8 日），https://securityintelligence.com/deeplocker-how-ai-can-power-a-stealthy- new-breed-of-malware/.

[23] *人工智能对网络安全的影响*：美国国家科学院、工程院和医学院研讨会论文集（2019），https://doi.org/10.17226/25488; 内克塔里亚 • 卡卢迪（Nektaria Kaloudi）和李静月（Jingyue Li），基于人工智能的网络威胁景观，ACM 计算调查，1-34（2020 年 2 月），https://dl.acm.org/doi/abs/10.1145/3372823；本 • 布坎南（Ben Buchanan）等人，《自动化网络攻击》，安全与新兴技术中心（2020 年 11 月），https://cset.georgetown.edu/research/ automating-cyber-attacks/；达科塔 • 卡里（Dakota Cary）和丹尼尔（Daniel Cebul），《破坏性网络操作与机器学习》，安全与新兴技术中心，5-23（2020 年 11 月），https://cset.georgetown. edu/research/destructive-cyber-operations-and-machine-learning/。

[24] *人工智能对网络安全的影响*：美国国家科学院、工程院和医学院研讨会论文集（2019），https://doi.org/10.17226/25488。

[25] 最近发生的 SolarWinds 攻击事件表明，我们的软件供应链存在着深刻的漏洞。见联邦调查局（FBI）、网络安全和基础设施安全局（CISA）和国家情报局长办公室（ODNI）、国家情报局长办公室的联合声明（2020 年 12 月 16 日），https://www.dni.gov/index.php/newsroom/press- releases/item/2175 -joint-statement-by-the-federal-bureau-of-investigation-fbi-the-cybersecurity-and- infrastructure-security-agency-cisa-and-the-office-of-the-director-of-national-intelligence-odni。

[26] *人工智能对网络安全的影响*：美国国家科学院、工程院和医学院研讨会论文集（2019），https://doi.org/10.17226/25488 本 • 布坎南（Ben Buchanan）等人，《自动化网络攻击》，安全与新兴技术中心（2020 年 11 月），第 3 期（2020 年 11 月），https://cset.georgetown.edu/research/automating-cyber-attacks/。

[27] 网络空间日晷委员会报告，美国网络空间日晷委员会（2020 年 3 月），https://www.solarium.gov/report。

[28] 参阅网络空间日晷委员会报告中的建议 5.4，第 87 页，美国网络空间日晷委员会（2020 年 3 月），https://www.solarium.gov/report。

[29] *对抗性人工智能威胁矩阵*：案例研究，GitHub（最后一次访问：2021 年 1 月 10 日），https://github.com/mitre/advmlthreatmatrix/blob/master/pages/case-studies-page.md。更多关于对抗性人工智能的应用，请参见纳维德 • 阿赫塔尔（Naveed Akhtar）和阿杰马勒 • 米安（Ajmal Mian）的*计算机视觉中深度学习的对抗性攻击的威胁*：调查，IEEE（2018 年 3 月 28 日），https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8294186。

[30] 对抗性人工智能是关于可以对人工智能系统做什么。保护和捍卫人工智能应用程序免受攻击的科学被称为“人工智能保障”。攻击人工智能每个技术组件的科学称为“反人工智能”。

[31] 拉姆 • 库马尔（Ram Shankar Siva Kumar）和安 • 约翰逊（Ann Johnson），机器学习系统的网络攻击比你想象的更为常见，微软安全（2020 年 10 月 22 日），https://www.microsoft.com/security/blog/2020/10/22/cyberattacks-against-machine-learning-systems-are-more-common-than- you-think/。

[32] 据估计，只有不到 1%的人工智能研发资金用于人工智能系统的安全。参见南森 • 斯特劳特（Nathan Strout），《人工智能的三大安全威胁》，安全与新兴技术中心（2019 年 9 月 10 日），https://cset.georgetown.edu/article/the-three-major-security-threats-to-ai/。

[33] 拉姆 • 库马尔（Ram Shankar Siva Kumar）等人，《对抗式机器学习-行业展望》，第 2 期，阿奇夫论文预印本网站（2020 年 5 月 21 日），https://arxiv.org/pdf/2002.05646.pdf。

[34] 公共和私人部门正在做出各种努力，例如，包括 MITRE-微软的对抗性机器学习框架。参见拉姆 • 库马尔（Ram Shankar Siva Kumar）和安 • 约翰逊（Ann Johnson），针对机器学习系统的网络攻击比你想象的更为常见，Microsoft Security（2020 年 10 月 22 日），https://www.microsoft.com/security/blog/2020/10/22/cyberattacks-against-machine-learning-systems- are-more-common-than-you-think/对抗性人工智能威胁矩阵：案例研究，MITRE（最后一次访问：2021 年 1 月 10 日），https://github.com/mitre/advmlthreatmatrix/blob/master/pages/case-studies-page.md。

[35] 类似的建议见密歇根大学 Michèle Flounoy 等人，《通过测试建立信任》，WestExec Advisors，第 27 期（2020 年 10 月）https://cset.georgetown.edu/wp-content/uploads/Building-Trust-T hrough -Testing.pdf.（弗卢努瓦（Flournoy）等人认为，“一个国家人工智能和 ML 红色团队作为测试对抗攻击的中心枢纽，将国防部操作员和分析师、人工智能研究人员、科研 [中央情报局（CIA）、国防情报局（DIA）、国家安全局（NSA）]以及其他适当的情报机构（IC）部分聚集在一起。这将是一个独立的红色团队组织，它将拥有技术和情报专业知识，在模拟的操作环境中模拟现实的对手攻击。”）

[36] 华大集团建立并运营中国国家基因库 - 中国政府的国家基因数据库。也是 COVID-19 测试的主要全球供应商，这有可能提供对大型国际遗传数据集的访问；截至 2020 年 6 月 30 日，其已经向包括美国在内的 180 个国家提供了 3500 多万个试剂盒，并在 18 个国家建立了 58 个测试实验室。请参阅柯丝蒂 • 尼达姆（Kirsty Needham）的特别报告：*COVID 为中国的基因巨头打开新的大门*，路透社（2020 年 8 月 5 日），https://www.reuters.com/article/us-health-coronavirus-bgi-specialreport/special-report-covid-opens- new-doors-for-chinas-gene-giant-idUSKCN2511CE。

[37] 参见 Richard Pérez-Peña，什么是诺维乔克--与纳瓦尔尼中毒事件有关的俄罗斯神经毒剂？《纽约时报》（2020 年 9 月 2 日），https://www.nytimes.com/2020/09/02/world/europe/novichok-skripal. html；2020 年遵守和遵守军备控制、不扩散和裁军协议和承诺（遵守情况报告），美国国务院 Pt. V（2020 年），https://2017-2021.state.gov/2020 -adherence-to-and-compliance-with-arms-control-nonproliferation- and-disarmament-agreements-and-commitments-compliance-report-2//index.html#_Toc43298166.

[38] 有关人工智能和生物技术关系的其他建议见本报告第 16 章。

第二章 未来防御的基础

人工智能使能未来防御

建造技术主干网络

加快采用现有数字技术

培训和教育作战人员

对下一代能力进行投资

普及推广人工智能的发展

自从冷战结束以来，美国军方一直拥有对于所有潜在对手的军事-技术优势。现在，它的技术受到了挑战，特别是来自中国与俄罗斯的挑战。高级军事领导人已经发出警告，如果目前的趋势线没有改变的话，美国军方将在来年失去它的军事-技术优势[1]。人工智能是挑战的关键性领域，因为我们的两个大国竞争对手相信利用人工智能系统和自主性，它们可以抵消我们的军事优势。在未来的几十年里，美国只有在指挥与控制、武器和物流领域加快采用人工智能传感器和系统，才能超越拥有尖端技术的对手。

国防部必须树立远大目标。到 2025 年，整个国防部人工智能广泛整合的基础必须到位。这些基础包括内部人工智能开发团队和关键行业合作伙伴都可以访问的通用数字基础设施、具有数字素养的员工队伍，以及能够提高效率的现代人工智能业务实践。这些基础是军事部门实现人工智能战备状态的先决条件，本报告第 3 章对此进行了讨论。

“到 2025 年，整个国防部人工智能广泛整合的基础必须到位。”

国防部在将人工智能等新技术和颠覆性技术融入军事行动方面远远落后于商业部门。不过，随着国防部在 2017 年启动 Maven 项目，形势开始出现好转。Maven 项目旨在通过识别无人机和其他平台系统捕捉的视频片段中的物体来简化情报分析人员的工作[2]。国防实验室和研究机构正在开展其他具有前景的创新活动，各军种正在进行概念验证和演示[3]。尽管如此，富有远见的技术人员和作战人员在很大程度上受困于过时的技术、繁琐的流程以及为过时的或竞争性目标设计的奖励机制[4]。因此，尽管体制存在问题，但是人们总能够采取变通方法和迂回策略获得成功。

> “……富有远见的技术人员和作战人员在很大程度上受困于过时的技术、繁琐的流程，以及为过时的或竞争性目标设计的奖励机制”

整合人工智能的障碍来自于许多方面。一直以来，国防部坚持以硬件为导向的工作方针，对舰船、飞机和坦克等硬件设施采取倾斜政策。不过，国防部目前正在积极推动从以硬件为导向的政府部门到软件密集型“企业”的转变。国防支出仍然集中在为工业时代和冷战设计的过时的平台系统上[5]。许多部门流程仍然严重依赖 PowerPoint 软件和手动完成的工作流程。目前，促进机器学习发展所需的数据仍然是分散的、凌乱的，甚至有时常常被丢弃。平台系统尚未实现互通互联。采购、研发和部署工作很大程度上不得不遵循一系列僵化的连续流程，这给人工智能的关键性早期持续实验和测试造成了障碍。即使前景看好的人工智能项目也未能如期交付，并且常常受到以营利为目的的商业企业提供的专业软件和数据存储服务的约束。部署人工智能应用的云基础设施的建设工作进展缓慢。在业界只需花费几个小时或数日即可完成的数据共享协议和软件升级常常延迟数月。各军种人员缺乏使用人工智能所需的技术培训和经验。

与此同时，国防部的官僚体制给国防部与技术公司之间建立合作伙伴关系，以及国家安全创新基地建设等关键性活动造成了阻碍[6]。更严重的是，官僚体制带来的错综复杂的混乱局面使得部分企业不愿与国防部开展合作，而经济上的不合理性甚至使许多初创企业也望而却步。在建造和整合用于人工智能战争的大型系统方面，传统的国防公司将继续发挥核心作用[7]。不过，即使对于拥有系统操控资源和专业经验的承包商而言，它们也遇到了流程上和技术上的障碍，不得不放缓人工智能系统建造和整合的步伐。

因此，变革绝非易事。它需要国防部长举全部门之力来加快采用新技术，需要新兴技术指导委员会推动相关工作的执行以及协调国防部和情报部门之间的优先事项。国防部长应指导下列五大领域内的行动：

建议

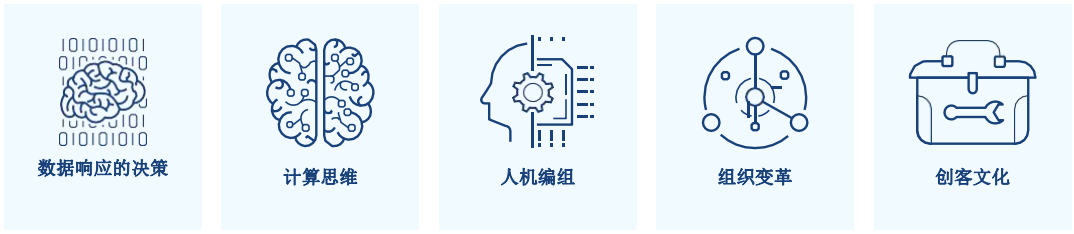
1. 建造技术主干网络。国防部应进行覆盖整个部门的基础性投资，用于支持无处不在的人工智能研发和部署的技术基础设施。2020 年，随着《国防部数据战略》的发布，国防部走出了充满希望的第一步[8]。不过，国防部缺少现代数字生态系统、协作工具和环境，以及对共享人工智能的广泛式随需访问——这些是在整个组织内部进行人工智能整合所需的必要元素[9]。国防部应当避免为每一个人工智能驱动的新项目或新能力进行重复性的核心基础设施改造，充分利用情报部门提供的成熟解决方案并与其进行互操作。覆盖整个部门的更广泛的平台系统，使得更加动态的开发和使用人工智能及更加高效地利用稀缺的专业技术经验变为可能[10]。

人工智能数字生态系统.

国防部长应当指导建立整个部门的数字生态系统，应当要求所有新的联合项目和服务项目必须遵循生态系统的设计，并且在尽可能的情况下，现有项目到 2025 年应与所述生态系统实现互操作[11]。关键要求应当包括：

- 数据架构由一个安全、联合的分布式存储库系统组成，该系统由数据目录和适当的访问控制链接[12]，有助于在整个部门查找、访问和移动所需的数据[13]。
- 打包的人工智能环境[14]，支持敏捷迭代的人工智能开发[15]、测试、部署和升级，以便向不同的利益相关者提供支持[16]。
- 共享型人工智能资源市场[17]，须建立在数据、软件和受训模型的联合存储库之上[18]；来自经过审查的云提供商资源池的预先协商的计算和存储服务。
- 加强型网络和通信骨干网，能够提供带宽来支持数据传输与融合、安全处理、人工智能应用程序的持续开发和部署，以及各级软件系统集成。
- 通用界面，允许迅速集成面向任务的投资。

“对于作战人员而言，思维方式决定作战方式。”



2. *培训和教育作战人员*。对于作战人员而言，思维方式决定作战方式。大多数现役军人利用性能强大的计算机来制作 PowerPoint 展示文件、电子表格和收发电子邮件。我们的军人需要提升核心能力，建造、使用并与机器系统进行编组，充分认识人工智能在构建高效作战部队方面的潜力。他们尤其需要了解：

- 在决策过程中，如何使用数据来提升人类直觉和经验。
- 如何使用信息处理代理以及如何让计算机执行人类无法有效完成的计算和分析。
- 如何在鼓励持续接触、定期试验和开发新工具的“创客”文化中发展和壮大。
- 如何转向与自主系统交互的“队友模型”并解决授权、可观察性、可预测性、可指导性和信任问题。
- 如何将组织带入人工智能时代，包括何时以及如何将人工智能相关工作整合到优先任务中、分配资源来构建和维护人工智能堆栈、监督新系统和支持技术专家的职业。

为了提升沿着上述路线进行的培训和教育质量，国防部应当：

- 在入职过程中，识别擅长计算思维的现役军人；
- 通过自我指导式教育课程和编码语言，对员工技能升级进行投资；
- 向初级领导者教授问题管理、人工智能生命周期、数据收集和管理、概率推理和数据可视化以及数据响应决策等课程，并将其作为任前要求和初始培训的一部分；
- 将新兴技术和颠覆性技术培训纳入专业军事教育课程；
- 创建新兴技术编码岗位和能够与联合岗位和认证体系相媲美的新兴技术认证项目。

建议

3. 加快采用现有数字技术。国防部在采用新技术时，很大程度上依靠变通办法，而它的核心采购流程仍然僵化。不过，我们注意到一些工作亮点，例如：发布可调整的采办流程、承包商资源[19]以及空军在某些项目上采取的做法[20]。国防部必须推广此类创新实践，并采取进一步措施来规范采办人员培训、项目奖励机制、预算和组织机构，以便为数字使能能力的交付提供更好的支持。

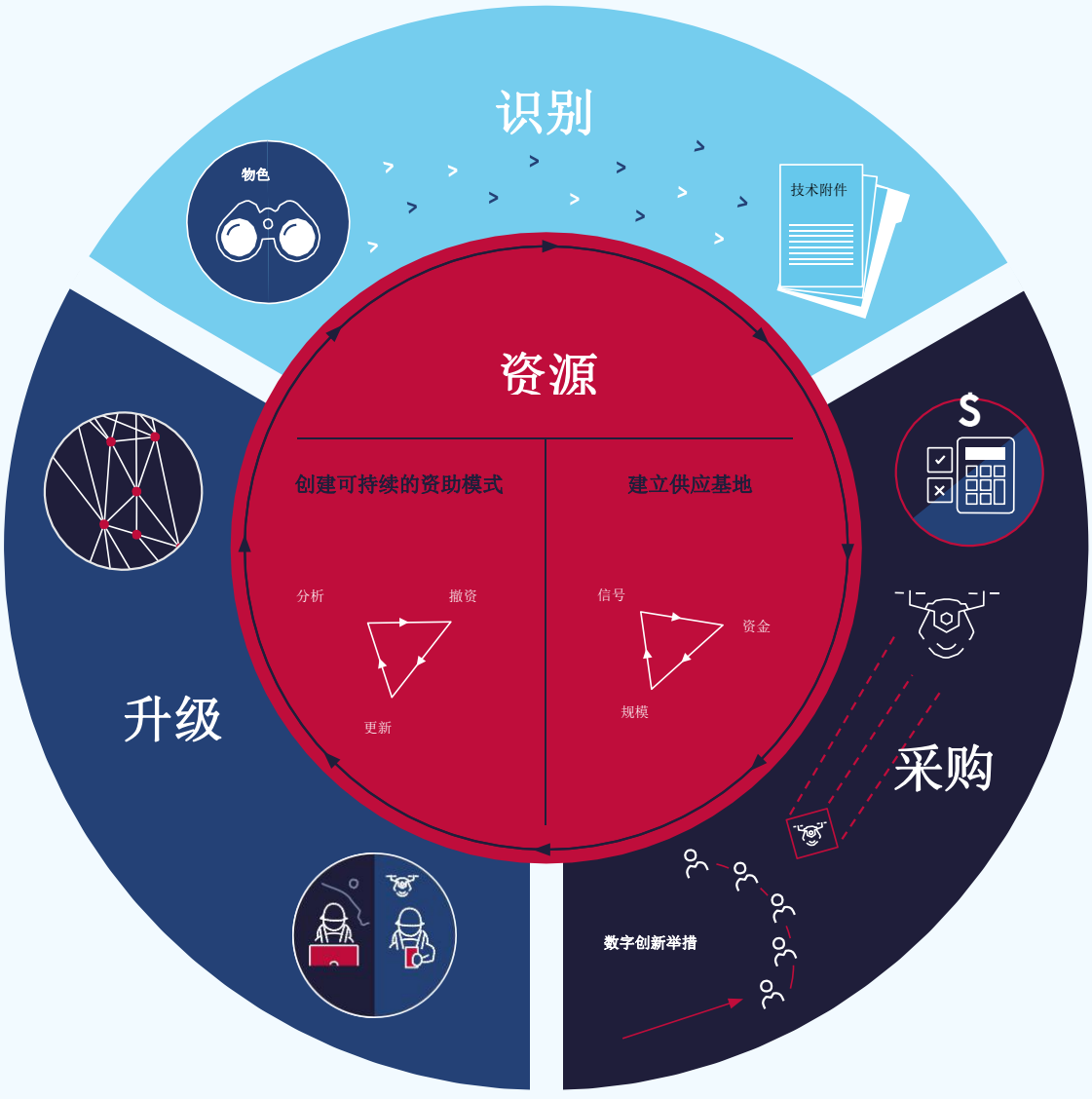
目前，国防部的许多数字创新项目已经实现交付[21]，但结果却不尽如人意，主要表现为缺乏协调和资源不足。国防部发出的“技术优先”的信号是临时性的，并且未获与非传统型供应商在数字技术领域进行重大投资业务记录的支持。因此，国家安全人工智能应用吸引的私人市场投资呈下降趋势。国防部应重点关注下列四项活动：

- **整合商业性人工智能，优化核心业务流程。**国防部应当接受成熟的商业性人工智能应用程序，鼓励使用相关程序来降低工作量和成本、推动实施各项行政措施并为决策提供信息[22]。作为关键性的第一步，国防部应当优先构建跨核心管理领域的企业数据集。
- **能够扩大影响的网络数字创新举措。**自下而上的创新需要与自上而下的领导相结合。国防部应当协调创新举措，对商业性技术解决方案实施统筹协调的上市战略。负责研究和工程的国防部副部长应与负责采办和保障的国防部副部长、各大军种和总部其他对应部门进行密切合作，为实施创新举措提供战略方向。
- **扩大使用专业的采办途径和承包方法。**国防部应当加快实施采办人才培训措施，培训内容包括所有可用的采办和承包选项，同时鼓励他们使用人工智能和数字技术[23]。

- **升级预算和监督流程。**国防部目前使用的资源分配流程与 1961 年实施的流程基本相同，因此与人工智能和其他数字技术并不兼容。国防部和国会应当采取一系列改革措施，通过加深理解速度、不确定性，实验和持续升级，推动软件和数字技术的进步。

高速推进人工智能的大规模交付

调整国防部和国家安全创新基地的流程、奖励机制和组织文化，使其成为对保持美国竞争优势至关重要的资源分配、优先处理、实施采购和迭代升级的一种综合性和战略性技术方法。



“在每一个层次上，技术人员、作战人员和领域专家都应组成一支综合性团队统一发挥职能。”

建议

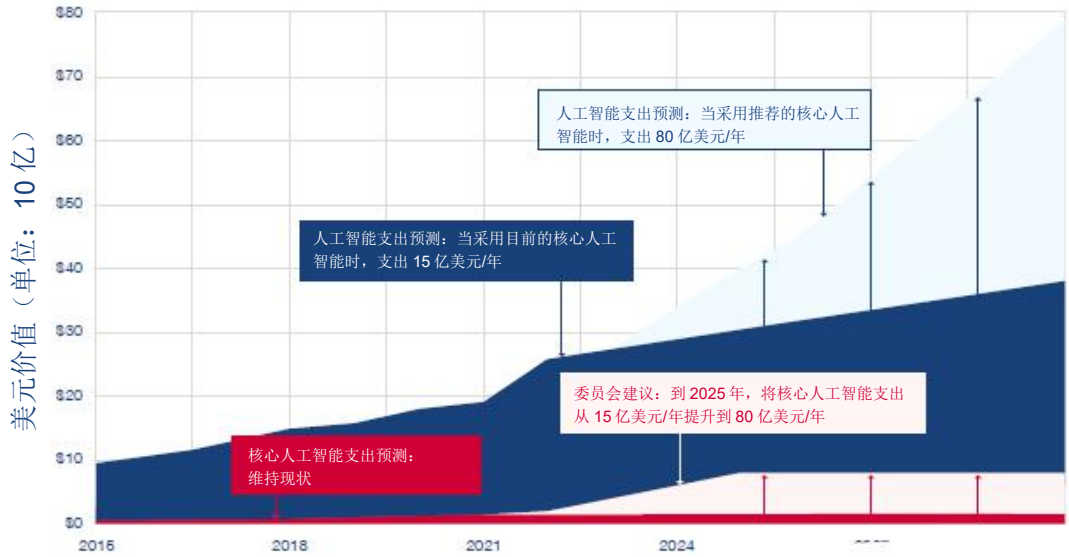
4. 普及推广人工智能的发展。国防部必须推动自下而上的人工智能发展[24]。在每一个级别上，技术人员、作战人员和领域专家都应组成一支综合性团队[25]。这将有助于用户进行反馈，并提升用户对人工智能系统的信任与信心。国防部应当：

- **指定联合人工智能中心作为本部门的人工智能加速器。**中心无法确立人工智能在部门内的所有作用，但是它能够并且应当可以发挥人工智能专业知识枢纽的作用。在“加速器”模型中，中心将与相关的采办、技术和管理办公室进行协调，为战略提供信息；有针对性地开发人工智能应用程序，解决作战司令部面临的共同挑战；为实现整个部门及所有军种的分布式人工智能发展提供资源[26]。

在 2015-2030 财年期间，加强人工智能研发投资

来源：戈维尼（Govini）和人工智能国家安全委员会

在 2015-2030 财年期间，加强人工智能研发投资



上图说明了核心人工智能技术研发投资与到 2030 年人工智能应用之间的相关性。图中展示了两种场景。在第一种场景下，国防部维持其当前的核心人工智能投资水平（每年约 15 亿美元）。在第二种场景下，国防部将核心人工智能投资水平提升到 80 亿美元/年。为了提高人工智能的应用率，需要大幅增加核心人工智能支出。

委员会工作人员与外部的两名合作伙伴共同分析了国防部过去在人工智能研发、测试和评估（RDT&E）上进行的投资以及计划进行的投资。分析数据来源于国防部年度 RDT&E 预算支出（2015-2020 财年）和年度 RDT&E 预算申请（2021-2025 财年）。有关所采用的方法和从这项工作中吸取的经验教训，请参阅本委员会存档的《美国国防部人工智能研发、测试和评估投资分析》。免责声明：我们认为该分析在人工智能支出总体趋势和解决方案方面产生了重要洞见，可以用于未来的更高水平分析。同时，我们提请阅读者注意：我们存档报告中详细阐述的源数据质量问题意味着所提供的支出水平估值包含重大且难以估计的误差幅度。

*人工智能*使能项目的开发（RDT&E 项目）和部署（采办项目）涵盖国防部所有作战系统和业务系统，集成了用于执行分析、自动化、通信、机动、监视、传感和其他多项任务的核心人工智能应用程序。尽管人工智能支出通常只占这些项目成本的一小部分，但是它们的系统性能严重依赖与核心人工智能的结合。

*人工智能*使能项目包括支持大规模部署人工智能能力所需的云计算和先进的微电子技术。

- **所有作战司令部都应组建综合性人工智能交付团队。**由于每个司令部都存在特定的军事行动要求，因此集中式开发无法满足实际需求。在这种情况下，每个司令部都应组建专属的人工智能交付团队来支持全生命周期的人工智能开发和部署，包括与通过数字生态系统利用公共资源相关的数据科学、工程、测试和生产[27]。交付团队还应包括可前出部署的分队，作为与作战单元的本地接口[28]。

5. 对下一代能力进行投资。国防部领导人预计未来的几年内，国防预算将呈现持平或下降趋势[29]。尽管存在潜在的预算压力，但是国防部必须通过优先考虑人工智能等新兴技术和颠覆性技术，继续加快推进现代化项目的实施[30]。

建议

- **资助人工智能研发。**国防部应当承诺每年将至少 3.4%的部门预算用于科学和技术，并且拨款至少 80 亿美元用于人工智能研发[31]。此外，国防部应将额外资源向具备显著人工智能专业能力的组织机构进行倾斜，例如：DARPA、海军研究办公室、空军科学研究办公室、陆军研究办公室和各军种实验室。

“国防部必须在预算编制时进行艰难的预算权衡，优先考虑现代化，从而实现人工智能在其业务流程和军事系统中无所不在。”

- **淘汰在人工智能战争中不具竞争力的老旧系统。**国防部必须进行艰难的预算权衡并优先考虑现代化，实现人工智能在其业务流程和军事系统中无所不在[32]。国防部应当寻求一种平衡方法，利用先进技术更新现有系统，为长期投资争取时间。此外，为了防止支持维护现状的偏见，国防部应当要求在向主要防御采办计划（MDAP）提供资金之前，对人工智能替代方案进行评估[33]。

- **制定国防战略的技术附件。**为了将国防部的技术投资战略与未来的作战需求联系起来，附件应包括设计、开发、部署和维持关键技术的路线图——这些关键技术对于解决国防战略中确定的军事行动挑战至关重要。

第二章- 尾注

[1] 时任参谋长、联席会议主席约瑟夫 • 邓福德（Joseph Dunford）将军在 2017 年证实：“美军对潜在对手的竞争优势正在削弱 .....我估计，五年内我们将失去部署力量的能力；这是我们保卫家园、推进美国利益和履行我们的联盟承诺的基础。”参谋长联席会议主席约瑟夫 • 邓福德（Joseph Dunford）将军在参议院军事委员会的姿态声明，参议院军事预算听证会第 2 页（2017 年 6 月 13 日），https://www.armed-services.senate.gov/imo/media/doc/Dunford_06 -13-17.pdf。

[2] *战争中的大数据：*特种作战部队，梅文计划（Project Maven），21 世纪战争，现代战争研究所（2020 年 8 月 25 日），https://mwi.usma.edu/big-data-at-war-special-operations- forces-project-maven-and-twenty-first-century-warfare/；谢丽尔 • 佩勒林（Cheryl Pellerin），梅文计划（Project Maven）在年底前将计算机算法部署到战区，国防部（2017 年 7 月 21 日），https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by- years-end/。梅文计划（Project Maven）现在包括检测、分类和跟踪全动态视频图像中的物体（如人员、车辆和武器）以及其他基于文本项目的人工智能算法。*PE 0305245D8Z：*情报能力与创新，国防部长办公室（2019 年 2 月），https://www.dacis.com/budget/budget_pdf/FY20/RDTE/D/0305245D8Z_187.pdf。

[3] 例如，陆军在 2020 年 9 月的“融合项目”演习中展示了在瞄准过程的多个阶段使用人工智能。珍 • 贾德森（Jen Judson）和南森 • 斯特劳特（Nathan Strout），在“聚合计划”中，美国陆军经历了成功和失败 - 它对这两者都感到高兴，《国防新闻》（2020 年 10 月 12 日），https://www.defensenews.com/digital-show-dailies/ausa/2020/10/12/at-project-convergence- the-us-army-experienced-success-and-failure-and-its-happy-about-both/.空军已经举行了类似的演习，最引人注目的是，作为其与高级战斗管理系统相关的努力的一部分 - 该技术基础设施将支持国防部的全域联合指挥和控制概念。特里萨 • 希钦斯（Theresa Hitchens），反导演示证明了 C2 的人工智能的能力，突破防御网（2020 年 9 月 3 日），https://breakingdefense.com/2020/09/abms-demo-proves-ai-chops-for-c2/。

[4] 这包括操作概念与技术发展相互作用的传统过程。本报告第 3 章提供了调整这种方法的建议，并确保技术进步既能为概念提供信息，又能推动技术开发。

[5] 一位观察员指出：“尽管国防部的投资账户在过去三年中大幅增长，但这一增长高度集中于从现有生产线购买系统和制作军用系统原型。”CSIS 国防工业倡议组主任安德鲁 • 亨特（Andrew Hunter）在美国众议院军事委员会上的证词，第 6 届国防部在与中国竞争中的作用听证会（2020 年 1 月 15 日），https://armedservices.house.gov/_cache/files/5/8/5818cc1f-b86f-4dca-8aee-10ca788e6f43/9F4A03ABF1DEAB747AF2D1302087A426.20200115 -hasc-andrew-hunter-statement-vfinal.pdf。

[6] 《国防战略》强调了国家安全创新基础在维护国防部技术优势方面的重要性。2018 年国防战略摘要，美国国防部，第 3 页，（2018 年），https://dod.defense.gov/Portals/1/Documents/ pubs/2018-National-Defense-Strategy-Summary.pdf。战略与国际研究中心提供了一个有用的定义，指出"[国家安全创新基地]与传统的国防工业基地概念相比，在范围上有很大的扩展[.....]"，包括硅谷、波士顿和奥斯汀等创新中心的科技公司。参见安德鲁 • 亨特（Andrew Hunter），《国防投资战略方针》，CSIS（2018 年 3 月 26 日），https://www.csis.org/analysis/strategic-approach-defense-investment。

[7] “国防部最大的六家主要供应商（洛克希德-马丁公司、波音公司、诺斯罗普-格鲁曼公司、雷神公司、通用动力公司和 BAE 系统公司）[...]占 2019 年国防部所有主要义务的 32%。”*2020 财年：*工业能力，美国国防部第 40 页（2020 年 12 月 23 日），https://www.businessdefense.gov/Portals/51/USA002573-20%20ICR_ 2020_Web.pdf?ver=o3D76uGwxcg0n0Yxvd5k- Q%3d%3d。

[8] 该战略为国防部将数据视为战略资产奠定了基础，并详细说明了使国防部数据可见、可访问、可理解、可链接、可信赖、可互操作和安全的目标。*执行摘要：*国防部数据战略，美国国防部（2020 年 9 月 30 日），https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF。

[9] 近年来，国防部已经采取了有希望的初步措施，为平台、云基础设施和软件开发建立管理服务结构。例如，空军的 CloudOne 和 PlatformOne 产品（https://software.af.mil/dsop/services/）；海军的黑珍珠 (https://blackpearl.us/)；以及陆军的编码存储库和转换环境（CReATE）。此外，国防部长办公室已经建立了一

个数据管理平台（ADVANA），目标是将其建立为审计和业务数据分析的单一权威来源。参见国防部副部长大卫 • 诺奎斯特（David L. Norquist）在美国参议院军事委员会战备小组委员会上的书面发言记录，6（2019 年 11 月 20 日） https://www. armed-services.senate.gov/imo/media/doc/Norquist_11-20 -19.pdf。

[10] 由于联合人工智能中心（JAIC）的联合共同基金会倡议，这个平台的组成部分正在进行，特别是共享人工智能资源的市场，包括数据、算法和训练有素的人工智能模型。

[11] 使用通用技术基础设施将会极大地提高国防部确保互操作性的能力，并提高联合部队的有效性。但必须指出，即使没有这样的关键技术基础设施，国防部正采取重要的政策步骤，推动旨在满足联合能力需求方案的互操作性和人工智能准备。参见 Aaron Mehta，Hyten 发布新的数据处理联合要求，国防新闻（2020 年 9 月 23 日），https://www.defensenews.com/pentagon/2020/09/23/hyten-to-issue-new-joint-requirements-on- handling-data/. 本报告第 3 章概述了关于在 2025 年前实现人工智能战备状态的其他建议。

[12] 对数据集和其他共享资源的安全访问应通过基于用户和角色的身份验证进行管理，并通过端到端身份、凭据和访问管理基础设施来实现。

[13] 这取决于国防部新数据战略的实施。*执行摘要：国防部数据策略*，美国国防部 （2020 年 9 月 30 日），https://media.defense.gov/2020/ DOD-DATA-STRATEGY.PDF。

[14] 这些平台环境具有现成的工作流，可以根据用户类型（如研究人员、行业伙伴、运营商）和使用情况（如开发、TEVV [测试、评估、验证和检定]、实地应用）进行定制和启动。

[15] 换言之，DevSecOps 应用程序生命周期。“DevSecOps 通过加强工程实践来改善交付结果的准备时间和频率，促进开发、安全和运营团队之间更有凝聚力的合作，因为他们致力于持续集成和交付。”理解敏捷和 DevSecOps 之间的区别-从商业角度看，GSA（上一次访问时间：https://tech.gsa.gov/guides/understanding_ differences_agile_devsecops/。

[16] 利益相关者可以包括在战术优势工作的嵌入式开发团队；提供预训练模型的私营部门合作伙伴；研究开放的相关挑战问题的学术研究人员；在服务实验室内工作的政府科技（S&T）研究人员；或共同开发可互操作人工智能能力的国际合作伙伴。

[17] 共享的人工智能资源应在持续的操作授权（ATO）框架下进行管理，并在整个部门中强制实行默认的 ATO 互惠。

[18] 类似于或依赖于 Git 的平台交付和特性(https://git-scm.com)，GitHub 网站(https://github.com)，和 GitLab(https://about.gitlab.com).

[19] 五角大楼采购办公室的“适应性采购框架”和“合同签订”标志着该部采取了重要步骤，以促进在采购和合同签订中使用替代授权。这些措施包括，例如，其他交易授权、中间层收购、快速原型设计和快速投入使用，以及软件收购的专门途径。

[20] 例如，空军的先进作战管理系统（ABMS），该系统作为一个组合来管理旨在支持新的全域联合指挥和控制概念的系统，并在很大程度上基于实验来推动创新和需求的迭代方法。值得注意的是，国防部 2021 财政年度拨款法案对空军方针的各个方面表示关注，包括“缺乏严格的要求、采购战略或成本估算”和系统集成系统。参见众议院报告 116-453，294-295（2020 年 7 月 16 日），https://www.congress.gov/116/crpt/hrpt453/CRPT-116hrpt453.pdf。

[21] 这里使用“数字创新倡议”一词来描述国防部长办公室和各军种的各种实体-- 国防创新组织 （DIU）、AFWERX、NavalX 和陆军应用实验室（AAL）--它们专注于弥合与商业技术部门的差距，特别是初创企业和非传统供应商，并加速交付最佳技术解决方案。

[22] 国防创新组织（DIU）目前正在实施一系列人工智能项目，以优化国防部的业务流程，从使用人工智能驱动的机器人流程自动化来减少陆军审计长的劳动成本，到使用人工智能驱动的预测性维护来改善空军的准备状态，以及利用人工智能构建的知识图谱来快速识别国防情报局的供应链风险。参见 JAIC Partners with DIU on AI/ML Models to Resolve Complex Financial Errors，JAIC（2020 年 10 月 1 日），https://www.ai.mil/blog_10_01__20_-jaic_partners_with_ diu_on_aiml_models_to_resolve_complex_finanical_errors.html；美国国防部授予 C3.ai 9500 万美元的合同车辆，以使用 ai 改善飞机准备状态，商业电讯（2020 年 1 月 15 日），https://www.businesswire.com/news/home/20200115005413/en/US-Defense-Department-Awards- C3.ai-95M-Contract-Vehicle-to-Improve-Aircraft-Readiness-Using-AI；Accelerate.AI 通过国防创新单位合同加速增长和产品采用，Accurate.AI（2020 年 4 月 23 日），https://blog. accrete.ai/newsroom/accrete.ai-wins-million-dollar-contract-with-the-defense-innovation-unit。

[23] 例如，国防创新组织（DIU）使用了多种获取途径和合同策略，这些途径和策略可以帮助提高人工智能解决方案的采用率和运营相关性，还可以扩展国家安全创新基地。国防创新组织（DIU）率先与新泽西州陆军合约司令部开创了商业解决方案，该司令部利用《美国法典》第 10 章第 2371B 节的其他交易授权，创建一个“快速、灵活和协作”的合约工具，为该门部提供原型能力。国防创新组织（DIU）还利用《美国法典》第 10 章第 2374a 节的奖项挑战权，为国防部和更广泛的人工智能研究界推进各种与人工智能有关的优先事项。

[24] 上述整个部门数字基础设施对推动这一方法至关重要，但也需要进行结构改革，以最大限度地发挥其效用。

[25] 国防部内有一些值得注意的战争技术人员配对的示例，例如空军的软件工厂和特种作战和陆军未来司令部使用的预先部署的战术数据小组。他们发现，技术人员（如数据科学家）与战术优势的操作人员或分析师结成合作关系：1）显著减少承包商了解问题集和部署解决方案所需的时间；2）激励迭代开发技术和快速推出最小可行产品，在加速的时间表上产生影响更大的解决方案；以及 3）产生更多对数据和人工智能技术的支持，作为关键的任务推动者。人工智能国家安全委员会义务（2020 年 11 月）。为了确保美军部队长期保持超强的战斗力，国防部必须扩大这种以用户为中心的发展。

[26] 与联合人工智能中心（JAIC）协调的重要办公室，包括但不限于负责研究和工程的国防部副部长、负责采购与保障的国防部副部长、作战行动测试与评估主管（DOT&E）以及国防部首席信息官（CIO）和首席数据官（CDO）。在负责研究和工程的国防部副部长职责范围内，国防创新组织（DIU）是联合人工智能中心（JAIC）的关键推动者，通过为特定应用程序转换商业原型来实现基于项目的方法。联合人工智能中心（JAIC）目前通过其部门任务倡议（CMI）为作战司令部提供服务，包括联合作战行动的任务倡议。见任务倡议，JAIC（最后访问日期：2020 年 12 月 28 日），https://www.ai.mil/mi_joint_warfighting_operations.html。

[27] 这种应用可以由其他作战司令部、军种软件厂或联合人工智能中心（JAIC）开发，并可通过推荐的数字生态系统发现。每个作战司令部应确保为人工智能交付团队配备适当的人才，以管理人工智能解决方案的整个生命周期，包括在数据科学、人工智能测试和模型培训、软件工程、产品管理和全栈开发等学科。

[28] 例如，陆军未来司令部（AFC）和美国陆军特种作战司令部（USASOC）都使用一种名为“战术数据小组”的模式。这种模式将人工智能/机器学习的专业知识以三到六人团队的形式带到现场，为实时运营问题建立人工智能解决方案。根据陆军未来司令部和美国陆军特种作战司令部的合同，由小型企业 Striveworks 执行，他们目前正在支持中央司令部和印太司令部责任区的工作。

[29] Jim Garamone，美国国防部讨论未来国防预算的主席（2020 年 12 月 3 日），https://www.defense.gov/Explore/News/Article/Article/2433856/chairman-discusses-future- defense-budgets/.

[30] *2018 年国防战略摘要，美国国防部，第 6 页，*(2018 年 6 月)，https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf。

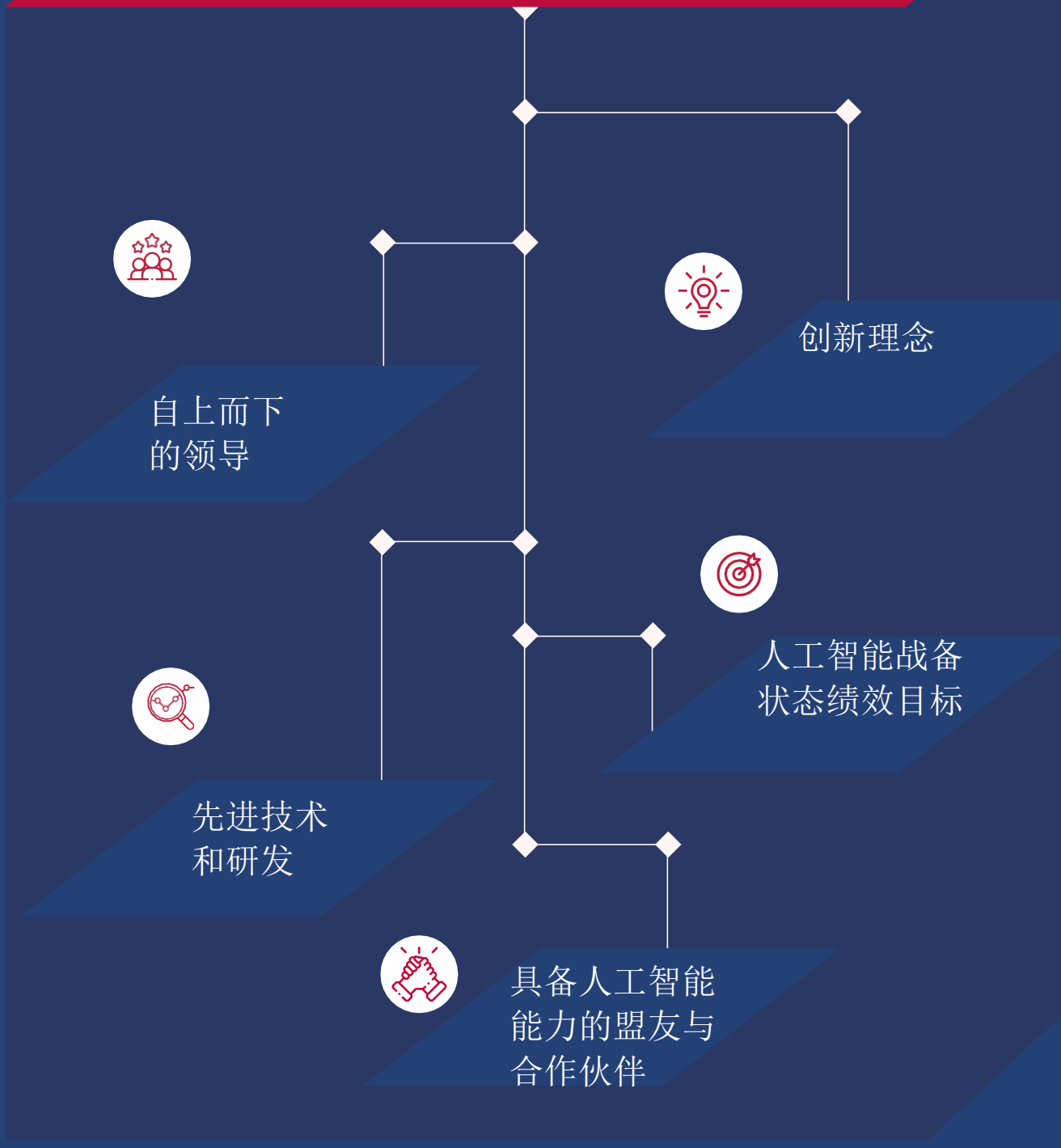
[31] 国防科学委员会过去曾提出 3.4%的水平，以反映私营部门的典型做法。*国防部研究、开发、测试和评估部（RDT&E）*：拨款结构，国会研究服务处，第 12 页（2020 年 10 月 7 日），https://fas.org/sgp/crs/natsec/R44711.pdf.

[32] 国防未来工作组的报告同样指出，“政策制定者、工业界和五角大楼必须共同努力，在国防机构内确定权衡，包括旧系统和行动，这将允许投资于技术和行动概念，以应对未来的挑战。”《2020 年国防工作队未来报告》，众议院军事委员会第 18 届会议（2020 年 9 月 23 日），https://armedservices.house.gov/_cache/files/2/6/26129500 -d208- 47ba-a9f7-25a8f82828b0/424EB2008281A3C79BA8C7EA71890AE9.future-of-defense-task-force-repor t.pdf。

[33] 这应该在可行的情况下利用战争游戏、实验和现场虚拟建设环境，并应规定与数字生态系统的互操作性。这一点与国防未来工作组相呼应，该工作组建议每个主要防御采购计划（MDAP）都应要求“在提供资金之前，至少评估一种人工智能或自主选择”。7.

第三章 人工智能和战争

到 2025 年，实现人工智能战备状态



即使实现人工智能战备状态的技术基础已经就位，但如果美国军方在整合人工智能技术时无法采用正确的理念、进行正确的操作，它将仍然在战场上处于不利地位。历史经验告诉我们，当新技术应用于军事用途时，是最好的应用者和集成者收获了新技术带来的丰厚军事成果[1]，而不是最好的技术人员。国防部不应成为人工智能革命应用于军事领域的见证者；相反，它应当通过自上而下强力领导、全新作战概念开发、持续不断作战实验和灵活性与风险并存的奖励体制，实现这一伟大变革。

得益于人工智能技术的发展，一种新型的作战模式正在形成。对此，我们的竞争对手进行了大量投资，希望占得先机。这种作战模式背后的理念被称作“算法战”或“马赛克战争”[2]；中国的理论学家将其称作“智能化战争”[3]。所有这些术语都以不同形式描述了人工智能将如何主导这个全新的冲突时代，以及在这个时代里，算法之间如何直接进行对抗。优势将取决于军方数据的数量和质量、军方开发的算法、连接的人工智能网络、部署的人工智能武器，以及为创造新的作战样式而采用的人工智能作战概念。

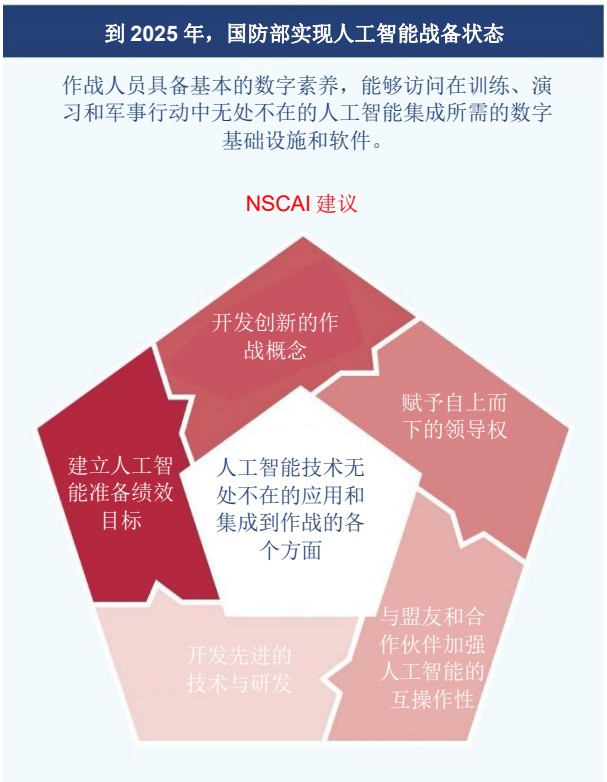
今天，国防部正在试图按部就班地按照计划创建人工智能枢纽，全然没有任何紧迫感。尽管国防部进行了一些富有想象力的改革并拥有一小批富有远见卓识的领导人，但是它仍然固守工业时代的思维模式，将大国之间的冲突看作是大规模作战部队及单一平台和系统之间的竞争。人工智能在商业领域的不断普及和数字化转型的速度凸显了无法进行快速调整的风险。国防部现在必须采取行动，在关键的功能、现有系统、演习和兵棋推演中实现人工智能集成，以便到 2025 年建立一支实现人工智能战备状态的武装部队。与此同时，国防部必须开发更具创造力的作战概念，实现作战概念与对未来人工智能技术投资之间的匹配与结合，从而不断超越潜在的对手。如果我们的军队未能配备以超越对手的全新作战概念为指导的人工智能系统，我们就将被战争的复杂性打败并且陷入瘫痪当中。

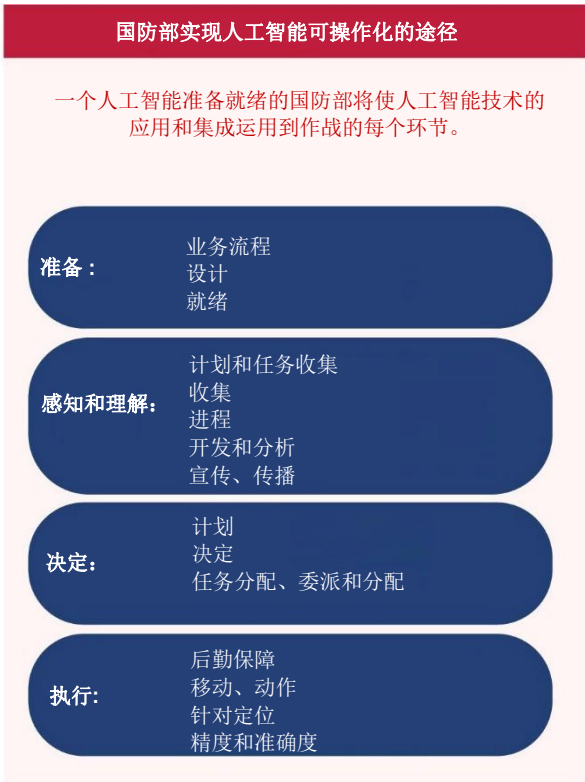
到 2025 年，实现国防部的人工智能战备状态：

作战人员具备基本的数字素养，能够访问在训练、演习和军事行动中无处不在的人工智能集成所需的数字基础设施和软件。

“国防部现在必须采取行动，在关键的功能、现有系统、演习和兵棋推演中实现人工智能集成，以便到 2025 年建立一支实现人工智能战备状态的武装部队。”

人工智能和战争





要在未来的冲突中与对手进行竞争、威慑对手，并在必要时与对手进行交战乃至取得最后的胜利，需要对作战概念、技术、组织结构及将盟友与合作伙伴融入军事行动的方式进行全面调整。与此同时，为了将面向未来的自主武器系统纳入作战体系，它还需要对广泛集成人工智能能力的利弊进行风险评估。最后，它还需要我们与盟友

和合作伙伴进行双边和多边对话，督促它们创建类似的人工智能枢纽，以此确保未来的互操作性。

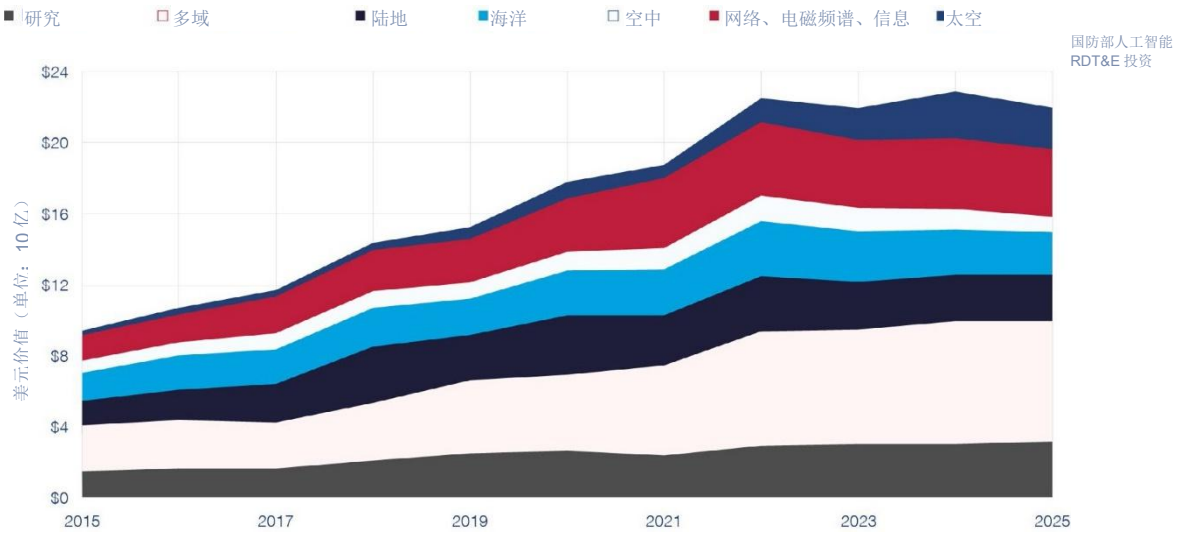
人工智能如何改变战争

基于人工智能的战争将不再取决于单一的新式武器、技术或作战概念；相反，它将聚焦于人工智能技术在作战各个方面的集成与应用。人工智能将改变从海底到太空、网络空间和电磁频谱等各个作战域的作战样式。它将影响战略决策、作战概念和规划、实地战术演练和后台支持。如本报告第 1 章所述，在这种新型战争中，通过人工智能支持的微小目标锁定、虚假信息和网络作战，传统的战场边界将被扩大。人工智能将重塑战争的多项属性，例如速度、节奏和规模；军方人员与智能机器之间的关系；对战场进行监控的持久性；以及对目标进行攻击时的辨别力和精确度。随着冲突的开展，发展知识、采取行动和做出反应的速度和准确性将更加重要。

“基于人工智能的战争将不再取决于单一的新式武器、技术或作战理念；相反，它将聚焦于人工智能技术在作战各个方面的集成与应用。”

按作战域划分的国防部人工智能 RDT&E 投资，2015-2025 财年

来源：戈维尼（Govini）



国防部人工智能投资广泛分布于不同的作战域（陆地、海洋、空中、太空、网络、电磁频谱、信息），其中超过 25%的投资用于人工智能的多领域应用，这显示了人工智能整合多领域作战的潜力。从 2019 财年到 2025 财年，国防部用于太空作战的人工智能应用投资增长了两倍以上，从 5 亿美元增至 22 亿美元，占人工智能投资的比例从 3%增至接近 9%。

注意：表中展示的支出水平是基于对国防部 2021-2025 财年 RDT&E 预算记录分析结果得出的估值。参见《国防部人工智能 RDT&E 投资分析》，人工智能国家安全委员会（最终在委员会进行存档备案）。由于源数据中固有的质量问题，表格中的估值包含重大且难以估计的误差幅度。

人工智能将使搜寻和打击具有军事价值的目标的速度更快、效率更高。它还将提升目标识别的准确率，最大限度地减少附带损害。目前，这一过程通常涉及以串行方式将一组数据从传感器传输给一系列工作人员，最后传输至可以对目标进行射击的平台。人工智能将有助于实现决策过程的部分中间阶段的自动化。人工智能还将为实施更加先进的流程创造机会。先进流程操作模式更类似于网络，能够将多台传感器和平台进行融合，共同管理复杂的数据流，并将行动信息传输至所有作战领域内的人类作战人员和智能机器[4]。

在战争中，人工智能的许多用途将补充而不是取代人类的作用。人工智能工具将改善军事人员在执行任务过程中，感知、理解、决策、适应和行动的方式。但是，全新军事行动概念也意味着作战方式不断发生变化，人类会将日益复杂的任务授权委托给人工智能系统来完成。在短期内，这将通过军方的“任务式指挥”原则进行管理，该原则强调下级应遵循指挥官的意图，以分权的方式执行上级的命令，并保持纪律严明的主动性。在可预见的未来，我们的武装部队仍将坚持这种以人为本的作战方式。随着人工智能继续在认知和神经形态领域不断深入，人机编组变得更加复杂，军方需要开发更具想象力的作战概念和组织结构，从而确保在不放弃“任务式指挥”原则的条件下，充分利用人工智能技术。

准备

业务流程：机器人流程自动化和人工智能使能分析能够显著节省成本、加速管理流程，向决策者提供关于财务、预算、合同、出行和人力资源等核心业务流程的卓越见解。

设计：人工智能将支持以成体系的“系统之系统”方法，通过数字工程、数字孪生、建模和仿真手段，进行兵力设计，更全面地了解体系脆弱性以及相关的能力、概念和技术。

战备：人工智能通过将重复性任务授权交付给能够更出色完成此类任务的机器，减轻从事重复性任务导致的认知负担，提升培训效果。人工智能将在所有演习和兵棋推演中普遍存在，增强武装部队在实际环境、虚拟环境和建设性环境下进行训练的能力。

规划和任务收集： 利用自动化技术，人工智能系统将近乎实时地分配和收集任务，从而满足动态情报要求或环境变化。

收集：在战术边缘，“智能”传感器能够预处理原始情报，明确传输和存储数据的优先级，这在恶劣环境或低带宽环境下将起到明显的作用。

处理：人工智能支持的自然语言处理、计算机视觉和视听分析能够大大减少人工数据处理。人工智能还可以用来进行数据自动化处理，例如翻译和解密等，从而加快获取有价值情报的能力。

感知与理解

利用与分析：人工智能使能工具可能增强跨多个数据集的过滤、标记和分类。此类工具能够以比人类分析人员更有效率和更大规模地识别联系和相关性，并对发现结果和最重要的内容进行标记识别以供人类分析。人工智能将提升对军事领导人的启示和警示。

- 人工智能可以融合来自不同来源、不同类型和不同分类水平的数据，并以目前不可能的方式产生准确的预测性分析。
- 目前，语音到文本转化和语义分析领域的技术进步能够支持阅读理解、知识问答和大量文本的自动摘要。

传播：人工智能将能够自动生成机器可读版本的情报并以机器速度进行传播。情报部门和军方的计算机系统能够在无需人工干涉的情况下，实时获取和利用这些情报。

决策

规划：人工智能决策将支持应用程序利用建模和仿真算法以及实时数据集来优化规划选项。

决策：人工智能将整合指挥和控制网络，提升定位和打击具有军事价值目标的速度。

任务分配、授权和分布：由授权部门加强的边缘处理将使前线作战单元在最低限度通信和无通信的情况下，以协调的方式进行军事行动。机器学习等人工智能技术和基于规则的模型将支持网络弹性。

后勤和保障：人工智能支持的预测性分析、优化和跟踪将提升后勤各个方面的效率和效果，有助于制定日常和紧急后勤保障行动方案。机器人自动化流程将简化以人为本的维护和供应链工作流。

执行

军事调动：人工智能将提升指挥官调动、驻扎和保护作战部队的能力。人工智能将通过人机编组和机机编组的方式来辅助作战网络及协调自主蜂群的行进。

目标：人工智能系统将把单一目标链扩展成考虑了跨军种和跨领域的多个变量的复杂目标网络。

精度和准度： 通过人工智能使能的智慧武器和自主平台，人工智能将使军方更精确地识别友军、非作战人员和敌方目标。

这份人工智能如何改变作战原则和能力的清单，以及其他类似的清单，绝非详尽无遗。创新将催生面向未来的能力。尽管目前我们对这些能力一所无知，但是随着时间的推移，它们的特质将逐渐显现并且愈发清晰。

> “创新将催生面向未来的能力。尽管目前我们对这些能力一所无知，但是随着时间的推移，它们的特质将逐渐显现并且愈发清晰”

携手合作，更加强大

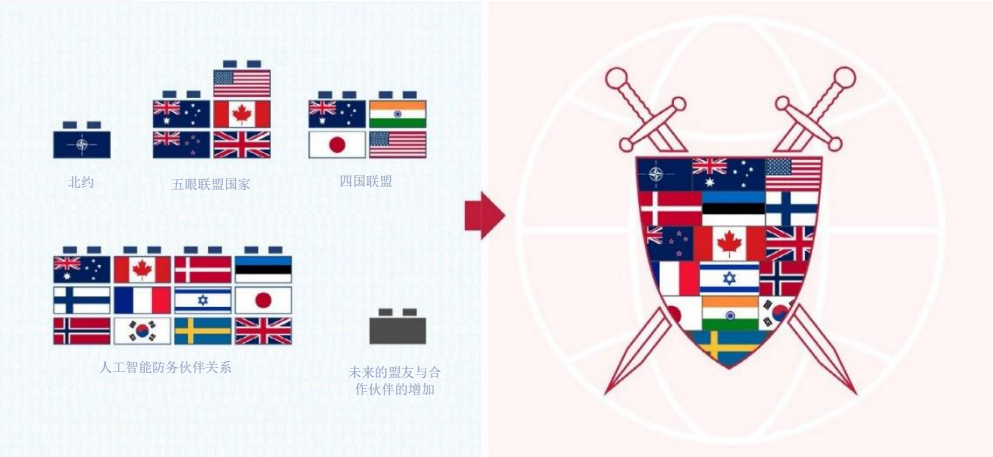
如果美国想在战斗中使用人工智能，它需要拥有具备人工智能能力的武装部队和情报部门的盟友与合作伙伴的支持。在应用人工智能方面，步调不一致将威胁美国加入的国际性和区域性联盟的互操作性、政治凝聚力和弹性[5]。随着美国在全球，特别是在欧洲和印度-太平洋地区，不断深化和扩大常规防御安排，我们应该将人工智能和新兴技术纳入综合性防御和情报活动。鉴于许多基于软件的能力具备双重用途，国防部需要更加灵活地与民间机构、企业、研究机构和盟国开展合作。

促进人工智能互操作性及关键性新兴技术在盟友与合作伙伴之间的普及，包括在五眼联盟、北约和印度-太平洋等国家和地区之间。具体措施应当包括：

- 加强五眼联盟成员国目前正在进行的与人工智能相关的防御项目和情报项目。
- 支持北约加速达成架构和标准协议，发展相关技术专长，并在演习和兵棋推演中寻求北约成员国的人工智能应用案例。
- 促进联合人工智能中心的国际人工智能防务伙伴关系，将其作为推进人工智能防务和安全合作的重要工具[6]。
- 建立大西洋-太平洋安全技术合作伙伴关系，提高欧洲和印度-太平洋地区的盟国及合作伙伴的军事和情报能力及互操作性。

“在应用人工智能方面，步调不一致将威胁美国加入的国际性和区域性联盟的互操作性、政治凝聚力和弹性。”

盟友与合作伙伴



到 2025 年，实现人工智能在军事领域的战备状态。

为了实现这一目的，国防部应当：

建议

通过自上而下的领导，推动组织改革。高级文职官员和军官应当确定明确的优先事项和方向，对下属进行授权，并在追求新技术过程中接受更高的不确定性和风险。具体而言，国防部应当：

- 成立高级别的新兴技术指导委员会，由国防部副部长、参谋长联席会议副主席和国家情报局副局长共同担任主席[7]；
- 确保具有重要作战经验的三星将军或海军将官出任联合人工智能中心主任，直接向国防部长或国防部副部长汇报工作；
- 任命负责研究和工程的国防部副部长为联合需求监督委员会的联合主席和首席科学顾问；
- 在每个作战司令部中，任命一名人工智能行动参谋来发挥类似军法参谋的作用。他/她应为人工智能专家，向指挥官和参谋提供关于人工智能系统能力和局限性的建议，识别人工智能系统的不当操作。

建议

开发融合了新型作战能力和新兴技术的创新作战概念[8]。作战概念力求实现跨军种和跨作战域的无缝互操作性。概念开发人员应与技术专家密切合作，共同制订武装部队在未来场景中最为有效的作战方式。他们应当假设在未来的战场上，人工智能能力将无处不在。此外，作战概念还可以推动未来的投资。

建议

到 2021 年末之前，建立人工智能和数字战备状态绩效指标[9]。为了在整个国防部实现更加实质性的人工智能整合，国防部副部长应当：

- 指导国防部各部门通过当前战备管理论坛和现有流程，评估军事人工智能战备状态。新兴技术指导委员会应当与负责人事与战备的国防部副部长和联合参谋部密切合作，确保建立的人工智能战备状态标准纳入各大军种的战备报告和资源分配战略。
- 指导各大军种加快审查人工智能领域的技术空白，提供招聘和人才管理策略[10]。
- 指导各大军种与负责采办与保障的国防部副部长、联合参谋部、国防后勤局和联合人工智能中心进行协调，尽可能地优先将人工智能集成到后勤保障系统。
- 将人工智能集成到主要的兵棋推演和演习中，以此推动技术应用的现场学习方法。作战人员需要在早期开发阶段就与人工智能能力进行持续交互，以便形成关于它们如何运作并影响任务的关键反馈。广泛的作战实验将加速概念开发，提升技术性能[11]。
- 通过作战实验室激励基金对开展的人工智能应用实验进行激励。基金可由新兴技术指导委员会进行监督[12]。

建议

到 2021 年底之前，定义联合作战网络架构。联合作战网络的关键目标在于建造一个安全、开放标准的系统网络，支持作战级别和跨域的人工智能应用程序集成[13]。它应当被各大军种访问和使用并且包含若干元素，例如：指挥控制网络；数据传输、存储和安全处理；武器系统集成。联合作战网络的技术基础设施应得到数字工程的最佳支持[14]。与此同时，网络应当能够与本报告第二章中描述的数字生态系统进行互操作[15]。

建议

投资能够支持未来军事能力的人工智能优先研发领域，包括：

研究领域	时间范围		关键挑战	类别
编组的未来	**短期目标**：为人机相互依赖关系的解读与设计提供支持。通过对相互依赖关系的正确管理，实现顺利高效的人机编组活动	**长期目标**：通过对不断变化的环境进行推理，采取灵活的编组策略，达到最佳编组效果。通过灵活的适应周边环境，获得编组弹性	理解相互依赖关系及动态适应对团队表现的影响	与人类一起工作
先进的场景解读	**短期目标**：感知操作环境的基本变化及提醒人类操作人员的能力，同时能够切换到一个更加适应和匹配当前环境的不间断模式（需要集成此类人工智能模块）	**长期目标**：维持在一系列复杂和动态的环境和场景下支持行动意识和洞见的不间断模式	纳入来自于复杂多变环境的多源和多模态信息	感知和感觉
智能边界设备，计算与网络	**短期目标**：边缘传感器中狭隘的人工智能应用，例如对高度对抗的作战空间进行监视的远程摄像头	**长期目标**：自主边缘设备能够进行动态学习、分享数据并与其他设备进行编组，同时进行情报数据收集、利用和保持；掌握域相关的物理操作	网络限制；尺寸、重量和功率（SWaP）	硬件，设备和机器人学
稳健且富有弹性的人工智能	**短期目标**：带有抗干扰和不可否认性的受训人工智能模型切换的标准实践	**长期目标**：通过加密技术等以隐私为中心的机器学习，在攻击面上具有弹性并能够安全学习的人工智能系统	多个攻击面；利用鲁棒学习应对敌方机器学习方法；应用安全技术的同时，保持高度的准确率	整合与保证
测试、评估、核查与认证（TEVV）	**短期目标**：人工智能测试、评估、核查与认证的共同框架	**长期目标**：对于利用动态学习、自我意识和进行监视操作的全自主人工智能系统进行测试、评估、核查与认证，自主人工智能测试范围应包括多组工作人员和智能机器	针对给定的使用案例，了解能够确定其可接受风险水平所需测试的数量和类型	
提供决策支持的集成式人工智能、建模和仿真	**短期目标**：高度受限场景和环境下的决策支持	**长期目标**：针对带有较长时间范围的开放式环境，提供实时决策支持及制定行动路线	多模态数据集成；评估仿真模型的预测保真度	学习与推理
自主人工智能系统	**短期目标**：在相对固定的可预测环境中，根据人类操作人员的授权指令，在较短时间内实施一定的自主操作，同时独立执行简单任务	**长期目标**：独立参与任务的时间更长，能够感知和理解动态变化的作战环境，同时执行涉及多智能体协作的复杂任务集。在这个过程中间，自主人工智能系统需要持续做出保证和进行自我监控。	在复杂、不断变化且不可预测的环境下，独立完成目标；了解与人类操作人员进行交流的方式与时间	
向更加通用的人工智能发展	**短期目标**：狭义人工智能的可解释性得到提升；实施迁移学习的方法；模型微调	**长期目标**：人工智能系统能够通过理解和处理作战环境进行学习，基于语境知识进行决策，积累经验性只是	揭开人类学习和推理的秘密；更全面的情境意识及问题解决能力	

第三章- 尾注

[1] 关于军事采用，请参阅迈克尔 • 霍洛维茨等的《*军事力量的扩散：国际政治的起因和后果*》，普林斯顿大学出版社（2010 年）。

[2] 国防部高级研究项目局（DARPA）马赛克战争的中心概念是围绕着“通过使用人类指挥和机器控制，快速组成和重新组合更多分散的美国军队，为美国军队提供适应性，为敌人提供复杂性或不确定性”。布莱恩 • 克拉克（Bryan Clark）等人，*马赛克战争：*利用人工智能和自主系统实施以决策为中心的操作，CSBA 第六届会议（2020 年 2 月 11 日），https://csbaonline.org/research/publications/mosaic-warfare-exploiting-artificial-intelligence-and- autonomous-systems-to-implement-decision-centric-operations/publication/1。

[3] 中国人民解放军已经制定了一个以人工智能为核心的所谓“智能化行动”的作战概念。在这一结构中，中国的理论是，在未来的冲突中，核心的较量将是对抗性的战斗网络，而不是传统的武器平台，信息优势和算法优势将成为胜利的决定因素。参见艾尔莎 • 卡尼亚（Elsa Kania），《中国人工智能军事创新》，CNAS 第 1 期（2019 年 6 月 7 日），https://www. cnas.org/publications/congressional-testimony/chinese-military-innovation-in-artificial-intelligence（美中经济与安全审查委员会的证词）。

[4] 参见《实时创建跨域杀伤网》，美国国防部高级研究项目局（2020 年 9 月 18 日），https://www.darpa.mil/news-events/2020 - 09 -18a。另请参见 *A 人工智能融合：*允许分布式人工智能增强多域操作和实时态势感知，卡内基梅隆大学（2020），http://www.cs.cmu.edu/~ai-fusion/overview。

[5] 关于与人工智能相关的军事互操作性挑战，请参阅艾瑞克 • 林-格林伯格（Erik Lin-Greenberg），*盟友和人工智能：运营和决策的障碍*，德州国家安全评论（2020 年春季），https://tnsr.org/2020/03/allies-and-artificial-intelligence-obstacles-to-operations-and-decision- making/.

[6] 政策的战略信息。”联合声明，人工智能国防伙伴关系（2020 年 9 月 15-16 日），https://www.ai.mil/docs/AI_PfD_ Joint_ Statement_09_16 _ 20.pdf。该伙伴关系在 2021 年 1 月举行了第二次正式对话。国防部联合人工智能中心促进第二次国际人工智能国防对话，JAIC（2021 年 1 月 27 日），https://www.ai.mil/news_01_ 27_ 21-dod_joint_ai_center_facilitates_second_international_ai_dialogue_for_defense. html。

[7] 委员会承认《2021 财政年度国防授权法》第 236 条，允许国防部长建立一个新兴技术和国家安全威胁指导委员会，该委员会由国防部副部长、参谋长联席会议副主席、负责情报和安全的国防部副部长、负责研究和工程的国防部副部长、负责人事和战备的国防部副部长、负责采购和维持的国防部副部长、首席信息官以及国防部部长认为合适的其他官员组成。但第 236 节中所述的结构并不包括来自情报局的领导，因此将无法推动预期的行动。参见出版授权号 116-283，威廉 • 索恩伯里（William M. (Mac) Thornberry）《2021 财政年度国防授权法》，134 法令 3388 (2021)。

[8] 值得注意的是，国防战略强调需要“发展创新的作战概念”和“培养实验和计算风险的文化”。概念作者和技术专家之间更紧密的协调，将会创造一个更有活力的技术开发和整合周期。2018 年国防战略摘要，美国国防部，第 7 页，（2018 年），https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense- Strategy-Summary.pdf。

[9] “战备”是衡量军事效力的一个关键标准，并且仍然是关于国防准备的预算、政策和监督辩论的核心。在这种情况下，国防部应建立关键的人工智能和数字准备的性能目标，以衡量和推动部门和服务的问责制。见 G。James Herrera，《军事准备的基本原理》，国会研究服务第 2 期（2020 年 10 月 2 日），https://fas.org/sgp/crs/natsec/R46559.pdf。

[10] 如本报告第 6 章所述，已经确定需要建立数字兵团、民用和军用人工智能和人工智能相关的职业领域，扩大招聘途径，并建立招聘办公室。军事部门需要评估这些领域和机构的人员数量，而不是建立这些机构的必要性。

[11] 虽然人工智能将在所有领域无处不在，但与空间、网络和信息操作领域相关的高数据量使得这些领域的使用案例特别适合在战争游戏、演习和实验中优先整合人工智能的应用。

[12] 作战实验室奖励基金旨在刺激实地实验和示范，以“评估、分析和提供对更有效地使用当前能力的洞察力，并确定将技术纳入未来行动和组织的新方法”。参见《美国国防部作战实验室激励基金和治理结构国防部副部长备忘录》（2016 年 5 月 6 日），https://defenseinnovationmarketplace.dtic.mil/ wp-content/uploads/2018/02/DSD_memo.pdf.

[13] 所设想的网络与国防部正在进行的努力非常一致，即接受标准驱动的互操作性、系统适应性和数据共享。参见美国国防部海军部长、陆军部长和空军部长关于服务采办执行官和项目执行官的备忘录（2019 年 1 月 7 日），https://www.dsp.dla.mil/ Portals/26/Documents/PolicyAndGuidance/Memo-Modular_Open_ Systems_ Approach.pdf.。

[14] 如国防部数字工程战略中概述的目标和重点领域；作为数字工程知识体系一部分的术语、任何知识和准则；以及纳入《2020 财政年度国防授权法案》第 231 条，该条要求建立数字工程能力，以实现测试和评估的自动化。请参阅 国防部数字工程战略、国防系统工程部副助理部长办公室（2018 年 6 月）；另请参阅出版授权号 116-92，《2020 财政年度国防授权法》，133 法令 1198 (2019)。有关数字工程知识体系的描述，请参阅 Andrew Monje，《美国国防部国防部基于模型工程的未来方向》（2020 年 1 月 27 日），https://ac.cto.mil/wp-content/uploads/2020/05/RAMS-Monje-27Jan2020 -Future.pdf.

[15] 有关该架构应如何与数字生态系统互动的详情，请参见第 2 章行动蓝图。

第四章 自主武器系统和与人工智能战争相关的风险

降低与人工智能武器系统相关的战略风险

继续严格的测试、评估、核查与认证（TEVV）程序

制定实践的国际标准

与竞争对手探讨风险问题

对特定的人工智能应用程序进行限制

目前，全球军事大国，无论大小，正在积极探索人工智能和自主武器系统。此类系统拥有巨大潜力，能够协助指挥人员以高速度、高质量和更具针对性的方式进行决策。人工智能使得武器系统在性能、速度和分辨力方面都远远超过了人类，并且能够完成迄今为止不可能完成的任务。如果设计、测试和使用得当，它们可以减少意外交战的风险，减少平民伤亡，最大限度地保护基础设施免受附带损害，支持对作战人员及其指挥链的决定和行动进行详细审查，从而更加符合《国际人道主义法》[1]。尽管美国的武器平台已经使用自主功能超过 80 年[2]，但是人工智能技术有潜力实现更为新颖复杂的自主攻击和防御能力。

随着人工智能技术越来越多地应用于武器系统，我们不得不面临一系列重要问题：此类系统是否合法、安全并符合道德。对人工智能技术应用于武器领域持批评意见者认为，世界各国应该通过协商的方式，对于此类系统及其用途进行约束或限制。同时，部分人士担心自主武器系统可能会使冲突升级的可能性大大增加。与此同时，在确保此类系统能够最大程度减少意外军事交战的可能性，或降低意外和无法控制的冲突升级风险的具体实施步骤上，辩论仍在继续。自从 2014 年以来，《联合国常规武器公约》政府专家小组多次召开缔约国会议，讨论在“致命自主武器系统时代出现的新兴技术”[3]的技术、军事、法律和伦理维度。具体而言， 政府专家组正在审查自主武器是否符合《国际人道主义法》，以及是否应采取额外措施来确保人类在使用自主武器上，保有适当控制权。

委员会与民间团体、学术组织和政府机构就围绕人工智能和自主武器系统的法律、道德和战略问题展开了一系列讨论，包括潜在的军事利益和风险、可能出现的道德问题、国际社会的监管努力以及对《国际人道主义法》的遵守程度。委员会对讨论的结果进行了总结与提炼，提出以下四点意见：

意见原则 1：在人类指挥官或作战人员的授权下，正确设计和测试的人工智能和自主武器系统可以继续以符合《国际人道主义法》的方式进行使用。

该意见基于《国际人道主义法》的若干要素：

- **区分：**这一原则主张武装冲突各方必须对平民和作战人员进行区分[4]。随着人工智能目标识别系统的准确率日益提高，武器系统有可能减少目标误认的情况发生（这是造成无意交战的主要原因），从而降低平民伤亡和附带的基础设施损害[5]。

人工智能国家安全委员会关于人工智能和自主武器系统的判断原则

- 在人类指挥官或作战人员的授权下，正确设计和测试的人工智能和自主武器系统可以继续以符合《国际人道主义法》的方式进行使用。
- 国防部现有程序能够确保美国部署安全可靠的人工智能和自主武器系统，并以符合《国际人道主义法》的方式对其进行使用。
- 几乎没有任何证据可以表明美国的竞争对手拥有同等严格的程序来确保负责任地设计以及合法使用人工智能和自主武器系统。
- 委员会不支持在全球范围内禁止人工智能和自主武器系统。

- **相称性：**这一原则禁止武装冲突各方发动平民生命损失超过预期军事收益的攻击[6]。人工智能和自主武器系统可以被设计为能够并且应该根据人类关于攻击相称性的判断和指令，采取相应的军事行动。权衡预期军事收益与潜在平民伤害所涉及的道德推理仍将是人类指挥官的责任[7]。

- **问责制：**确保问责制和指挥责任对于遵守《国际人道主义法》至关重要。人类能够并且应该对于自主武器系统、人工智能武器系统和其他武器系统的开发、测试、使用和行为承担相应的责任。自主武器系统在与人类指挥和控制系统相同的一般参数内实施操作，后者是为了确保行动问责制落实到位及遵守《国际人道主义法》而专门设计的。在这一点上，所有武器系统并无差别[8]。

国防部下属的政策机构认为，在武装冲突中，当交战各方决定采取能够夺去人的生命的行动时，必须考虑到人类的判断。对此，委员会表示赞同和支持。由于实际情形时间关键性、作战环境和武器系统类别的不同，人类对于使用自主武器系统应承当的责任也相应存在差异[9]。世界各国有义务通过建立流程，确保在使用人工智能和自主

武器系统的过程中，依靠适当的人类判断，同时确保人类作战人员对使用此类武器系统的后果承担责任。

> “在武装冲突中，当交战各方决定采取能够夺去人的生命的行动时，必须考虑到人类的判断。”

人类对于致命性军事行动的后果承担责任，并不意味着人类需要监视交战过程的每一步。一旦人类授权人工智能武器系统或自主武器系统对单一目标或一组目标发动攻击，攻击过程中的后续步骤可以通过完全自主方式进行，但是人类仍然需要对后续步骤产生的结果负责。在上述攻击行动序列中，准确的步骤数取决于系统的技术性能和所在环境，并且必须考虑系统行为和潜在后果的不确定性、威胁大小及攻击行动的时间因素。举例来说，长时间处于快速变化的环境（如城市环境）中的自主武器系统，与运行几乎相同时长、但处于水下或太空等高度可预测的少人环境中的同等武器系统相比，可能需要更加频繁的人类授权来确保人类对自主行动的后果承担充分责任。这一逻辑可以并且应该纳入武器系统的设计、测试和作战规划。考虑到这些因素，在策划具备可行性和必要性的军事行动时，应将一组自动化操作中的人工指导因素纳入考虑范围。在这种情况下，人类必须对系统的状态进行检查并对下一步行动进行授权，然后系统才能继续履行使命。在涉及致命武力的交战过程中，强迫要求每个独立行动步骤都必须获得清晰的人类授权的一揽子计划既不现实，也不必要。事实上，这一政策反而可能促使指挥官采用不太精确的非制导武器系统，并因此造成更大程度的附带伤害。

意见 2：国防部现有程序能够确保美国部署安全可靠的人工智能和自主武器系统，并以符合《国际人道主义法》的方式对其进行使用。

国防部承诺，遵守严格的自主武器系统开发和使用程序及强有力的人工智能道德伦理原则。这一姿态让公众对国防部能够部署并合法使用人工智能和自主武器系统充满信心。目前，国防已经建立了综合性流程来确保在使用部署的任何武器过程中，遵守《国际人道主义法》，并且公开承诺在《国际人道主义法》的框架下进行军事行动，以此最大程度降低平民伤亡并从错误中吸取教训[10]。为此，国防部组建了跨部门的战争法工作组来“制定和协调战争法行动和问题，例如对国防部各部门正在考虑的新型作战手段或方法进行分析”[11]。这一常设机构有能力研究科技发展对《国际人道主义法》的

影响。红十字国际委员会高度赞扬了这一制度的活力和透明性，并将美国列为“拥有审查武器合法性的国家机制及向红十字国际委员会提供机制建立文书”的 8 个国家之一[12]。

除了基线法律审查以外，国防部还采取了自主武器系统特别预防措施，确保此类系统经过充分的测试、评估、核查和认证（TEVV）。2012 年，国防部发布了第 3000.09 号政令——《武器系统中的自主性》。这份新出台的武器开发指导性政令为国防部开发和使用自主武器系统提供了政策指南。它要求所有的自主武器系统在设计时，“必须允许指挥官和作战人员在使用武力问题上，能够给予适当水平的人类判断”，以及“在开发和部署致命自主武器之前，必须分别得到国防部高级领导人的批准”[13]。它还硬性规定，任何自主和半自主武器，在其操作状态修改后，必须再次进行测试与评估。国防部在 3000.09 号政令中对自主武器的重要定义和基线要求进行了详细阐述。与此同时，国防部必须每年对政令的适用性进行审查，确保政令与技术进步相匹配[14]。报告第 7 章对于美国应该如何通过调整 TEVV 政策和能力来确保对人工智能系统建立合理的信心给出了具体建议[15]。

“美国遵守《国际人道主义法》的承诺由来已久，人工智能和自主武器系统不会改变这一点。”

此外，国防部在指挥和控制程序中，对授权进行目标选择和使用弹药做了严格规定，确保遵守《国际人道主义法》。各级律师直接向战地作战指挥官提供支持，就使用武力的决定提供咨询意见。美国遵守《国际人道主义法》的承诺长由来已久，人工智能和自主武器系统不会改变这一点[16]。相同的原则将根植于武器设计中，显露于 TEVV 工作中，并且由监督武器部署的指挥官进行守护和捍卫。国防部在《武器系统自主性政策》及《2020 人工智能道德原则》中进一步强调了这一承诺[17]。

意见 3：几乎没有任何证据可以表明美国的竞争对手拥有同等严格的程序来确保负责任地设计以及合法使用人工智能和自主武器系统。

战场上的成功可能越来越依赖于人工智能系统的性能。鉴于人工智能的开源性和双重用途，可能出现人工智能武器扩散现象，并使世界各国面临迅速部署未经测试的新型系统和算法的压力不断增加——这种压力促使各国越来越倾向于设计能够迅速进行反应的武器系统，从而减少了人类对于交战决策进行有效监督的时间。与此同时，美国的竞争对手，特别是俄罗斯与中国，可能没有同等的操作和定位程序来确保以遵守《国际人道主义法》的方式来使用此类系统，或对使用致命武力采取问责制度。迄今为止，在采购、开发、测试和部署自主武器系统的政策和流程上，俄罗斯与中国没有公开出版或发行过任何相当于 3000.09 号政令的文件。与美国不同，在俄罗斯和中国，这些流程属于高度机密——当然，前提是这些流程的确存在。

美国的竞争对手已经表明它们在开发和使用人工智能武器系统时，不可能遵守相同的道德原则和法律标准。特别是俄罗斯，这个国家历来愿意部署有风险的、未经测试的武器系统。举例来说，它在叙利亚的战斗中部署了性能不佳、自主功能有限的无人地面车辆[18]。

“缔结一份禁止开发、部署或使用人工智能和自主武器系统的国际公约，不符合美国的当前利益，也无助于国家安全局势的稳定……”

意见 4：委员会不支持在全球范围内禁止人工智能和自主武器系统。

缔结一份禁止开发、部署或使用人工智能和自主武器系统的国际公约，不符合美国的当前利益，也无助于国家安全局势的稳定。基于下列原因，委员会不建议美国政府进行此方面的尝试：

- 首先是基本的定义问题。关于自主武器系统，虽然联合国早在 2014 年就已经主持进行了关于致命自主武器系统的讨论，但是截至目前，世界各国仍然未能就定义问题达成一致。在这种情况下，进行相关谈判存在很多问题，因为我们无法在不过度限制美国现有的军事作战能力的同时，对亟待限制的系统类别给出清晰明确的定义。

- 即使定义问题能够克服，我们认为在当前阶段，执行这样一份国际公约很不现实，原因在于无法对公约的遵守程度进行核查。目前，世界各国无法通过可行的技术方式相互证明特定的武器系统是否具有自主性，或是否具备或缺少特定的作战能力。要实现上述目的，一个国家必须允许外国检查人员在短时间内访问此类武器系统的底层代码，而世界各国不可能同意这样一种入侵式核查措施，因为披露相关信息意味着它们的系统安全将面临无法接受的风险。

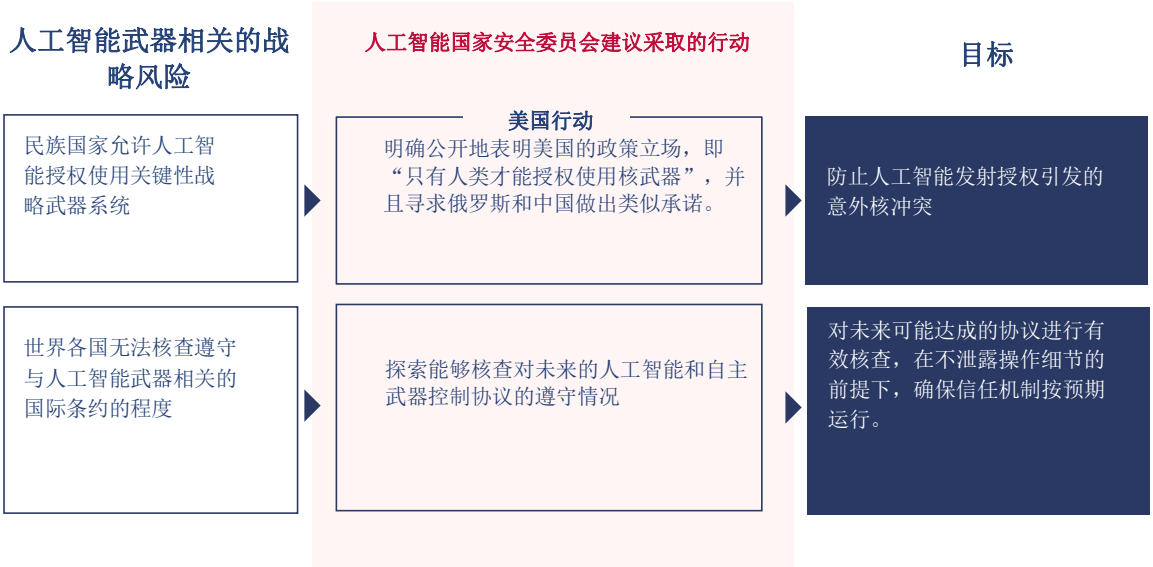
- 此外，国际公约造成的影响很可能与美国的战略利益背道而驰。俄罗斯或中国等国家可能做出空洞的承诺。因此，对于最有可能以不安全和引发一系列道德问题的方式部署自主武器系统的国家而言，缔结国际公约不会给其带来任何政治压力。相反，遵守国际法的国家，包括美国及其实行民主制度的盟友与合作伙伴，将首当其冲受到公约带来的压力影响。而且，由于美国盟友之间对于禁止自主武器系统持有不同观点，缔结国际公约将造成这些国家在利用人工智能支持的自主武器系统问题上的立场分裂加剧。如果美国的盟友加入国际公约而美国选择无动于衷，这种分歧将阻碍彼此之间的军事互操作性[19]。

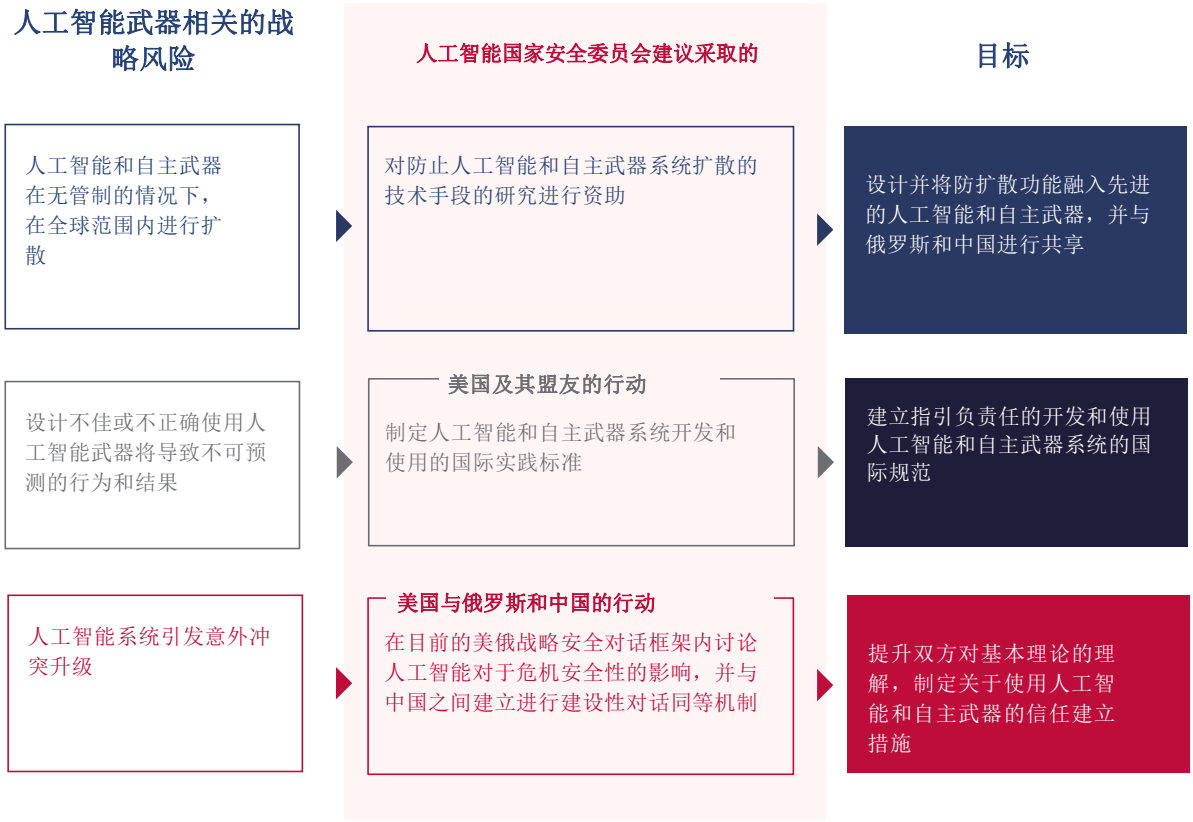
基于上述原因，我们认为国际公约带来的现实性战略问题远远超过美国及其盟友与合作伙伴可能获得的潜在利益。因此，我们支持美国政府在缔结国际公约问题上所持的反对立场。尽管如此，这并不排除通过其他协议或政策解决人工智能和自主武器系统带来的风险，也不排除当公约遵守情况可以核查时，未来对人工智能武器和自主武器特定技术的类别进行规定的可能性。

降低人工智能战略风险的建议

尽管委员会认为正确设计、测试和使用的人工智能和自主武器系统将带来重大的军事效益甚至人道主义利益，未受控制的全球性使用此类系统将会带来意外冲突升级的风险和危机不稳定性。美国不能假设其他国家部署的人工智能和自主武器系统，在开发、获取和部署阶段，经历了正确的测试与核查程序并且能够满足设计初衷。意外冲突升级可能存在很多原因，例如由于战场上部署的人工智能和自主武器系统之间交互的挑战性和未经测试的复杂性，以及在更普遍的情况下，智能机器或人类对信号或行动的错误解读，导致系统无法按预期运行。人工智能系统可能会全面提高战争的速度和自动化程度，减少冲突降级的时间和空间。我们不能假设其他国家开发的人工智能和自主武器系统会遵守《国际人道主义法》。

因此，世界各国在讨论人工智能和自主武器系统的开发、部署和使用时，必须采取行动，着眼于降低人工智能和自主武器系统的风险，提倡安全性，遵守《国际人道主义法》。鉴于美国的技术经验、军事实力以及在部署和使用人工智能和自主武器系统方面清晰透明的政策和道德准则，美国应该而且必须领导上述行动。针对如何降低人工智能和自主武器系统的相关风险，委员会建议美国政府采取下列五项行动：





明确公开地表明美国的政策立场，即“只有人类才能授权使用核武器”，并且寻求俄罗斯和中国做出类似承诺。美国应该明确公开地发表声明，强调只有人类，而不是人工智能或自主系统，才能够决定是否授权使用核武器，并将美国政府的公开表态纳入国防部的下一份《核态势审查》[20]。声明将巩固和凸显美国当前政策，即“使用核武器的决定需要获得美国总统清晰明确的授权”[21]。它还显示了美国切实致力于以负责任的态度使用人工智能和自主功能，限制不负责任的能力，防止人工智能系统以危险的方式推动冲突升级。由于声明减少了竞争对手对于美国突然实施意外打击的恐惧，并对做出类似承诺的其他国家进行奖励，它还将产生稳定效应。

美国应该积极敦促俄罗斯和中国以及其他拥有核武器的国家发表类似的声明。虽然“只有人类才能授权使用核武器”的联合政治承诺无法得到确认与核查，但是这份声明能够稳定人心，并且对于典型的“囚徒困境”给出答案：只要各国相信其他国家不会建立有可能无意中引发大规模核升级且具有风险的指挥和控制结构，它们开发这种系统的动力就会减弱[22]。尽管这一规范在美国被广泛接受，但是俄罗斯与中国是否具有相同的战略关切还不得而知。

“世界各国在讨论人工智能和自主武器系统的开发、部署和使用时，必须采取行动，着眼于降低人工智能和自主武器系统的风险，提倡安全性，遵守《国际人道主义法》。鉴于……，美国应该而且必须领导上述行动。”

公共报告显示，俄罗斯此前曾经安装了能够自动进行核武器发射授权的“死亡之手”系统[23]，而中国代表在与美方代表进行的第二阶段对话中，对于要求中国做出同等承诺的要求表现的迟疑不决。如果俄罗斯和中国都不同意这样的建议，美国在国际社会发动和引导声势浩大的国际舆论压力，对俄罗斯和中国进行强烈谴责，并且强调俄罗斯和中国如何拒绝就以负责任的态度使用人工智能系统进行承诺。

建议

*在目前的美俄战略安全对话框架内讨论人工智能对于危机安全性的影响，并与中国之间建立进行建设性对话同等机制。*美国国务院与国防部应该在目前的美俄战略安全对话框架内讨论人工智能对于危机安全性的影响，并与中国之间建立进行建设性对话的同等机制。战略安全对话是一项机构间的双边对话机制，着眼于减少双方在关键的战略性问题和威胁上的误解，降低意外冲突升级的可能性。尽管传统上对话侧重核军备控制和理论，但是最近它也被用来探讨新兴技术和太空安全[24]。目前，由于中国近10年来一直拒绝美国在相关问题上进行对话的要求，美国与中国之间尚未建立同等的对话机制。尽管如此，在过去的一年里，越来越多的证据显示中国对于与美国在人工智能军事系统上开展正式对话很感兴趣[25]。美国应该培养和利用这种兴趣来建立包括双方军事、外交和安全官员在内的美中战略安全对话机制。

鉴于美国、俄罗斯和中国都在积极寻求人工智能能力，以及俄罗斯和中国有可能在部署人工智能系统时未经过严格的 TEVV 程序，并因此导致人工智能系统可能存在不安全或不可靠因素，三国应该增进对彼此军事理论的理解，包括对人工智能和自主系统的了解，这一点至关重要。美国应该利用这一渠道，强调部署不安全的系统可能导致意外冲突升级，指出进行严格的 TEVV 工作程序的必要性，并与俄罗斯和中国讨论如何看待常规冲突迅速升级的风险，从而更好地预测未来的危机应对措施。

“……应该增进对彼此军事理论的理解，包括对人工智能和自主系统的了解，这一点至关重要。”

这些对话还将播下未来的种子。在长期对话中，三国交流的重点应该集中在如何围绕人工智能和自主武器系统，制定切实具体的信任建立措施。举例来说，美国、俄罗斯和中国可以效仿 1972 年的《美苏防止海上事件协定》，共同起草一份“国际自主事件协定”，寻求为自主军事系统设定“道路规则”，以便创建更可预测的作战环境，避免事故和误判[26]。三国还可以同意在系统中集成“冲突自动升级绊线”，防止在无人干预的情况下，冲突在特定情境中自动升级——包括上述核武器的使用。

与盟友共同制定人工智能和自主武器系统开发、测试和使用的国际实践标准。美国必须与盟友开展密切合作，共同制定世界各国以负责任的态度进行人工智能和自主武器系统开发、测试与应用实践的相关标准。这项工作可以在目前取得的一系列工作成果基础之上进行，包括致命人工智能和自主武器政府专家组在 2019 年商定的 11 项指导原则[27]、国防部 3000.09 号政令、国防部《人工智能道德准则》以及人工智能国家安全委员会的《以负责任的态度开发和部署人工智能系统时的关键考量因素》[28]。作为上述举措的一部分，国防部战争法工作组应该定期召开会议，审查与自主武器系统和《国际人道主义法》相关的未来技术发展；新兴技术指导委员会（委员会在本报告的第 3 章中就此单独提出过意见）应该就未来技术发展对政策和国家安全的影响给出建议。

上述两个小组的工作成果，可以为未来国防部与盟国和竞争对手在人工智能和自主武器系统方面的合作提供参考。在此类系统的开发、测试和使用标准方面获得盟友的共识，将为这些系统制定重要的规范，有助于确保系统开发和使用的安全性，并进一步凸显美国及其盟友对以符合道德的方式和负责任的态度使用人工智能的承诺。美国还应该利用上述两个专家顾问团队的工作成果，强调人工智能将成为未来军事行动关键组成部分，并制定共同框架来指导以正确的方式和负责任的态度在战场上使用人工智能和自主武器系统。此外，美国还应积极鼓励盟友在本国部队的数字化和现代化方面进行投资，同时也应重点强调：如果任何盟友同意加入禁止致命性自主武器系统的条约，那么彼此之间的军事互操作性将面临风险。

建议

*寻求各种技术手段，对与人工智能武器系统相关的未来军控协议的遵守情况进行核查。*美国应该积极寻求开发相应的技术并制定相应的战略，以便能够安全有效地核查涉及使用人工智能技术的未来军控协议。尽管目前对人工智能武器系统的军备控制在技术上无法验证，但是进行有效的核查，对于在未来形成对人工智能能力的法律约束力，仍然很有必要。国防部和能源部应该带头设计和实施各种技术——这些技术可以让其他国家相信，人工智能和自主武器系统正在按计划工作，不会泄露敏感的操作细节。举例来说，美国可以研究如何让人工智能武器平台生成可验证的操作记录；当怀疑有不符合规定的活动时，可以通过国际社会开展的质疑检查活动对其进行现场检查。因此，美国有必要进行技术创新，以便在不泄露敏感信息的情况下，对人工智能能力在全球范围内进行限制。

建议

*对能够防止人工智能和自主武器系统扩散的研究进行资助。*鉴于人工智能算法的开源性、双重用途和固有的可传播性，控制人工智能和自主武器系统扩散带来重大挑战[29]。主要利用商业组件的临时自主武器系统的扩散很难通过监管进行控制。在这种情况下，我们应该积极开展情报共享和国内执法工作，防止此类武器落入恐怖分子和其他非国家行为体的手中。关于更复杂的自主武器系统，美国应该加倍努力设计和纳入防扩散功能，如防止未经授权的用户使用此类武器的标准化方法，或通过改变关键系统参数对系统功能重新进行编程。国防部和能源部应该资助相关的技术研究，并在适当条件下，与俄罗斯和中国，或潜在的其他国家进行分享，共同防止某些人工智能自主武器系统的扩散或失控[30]。

本报告不包含第 4 章的单独行动蓝图。这是因为，鉴于该主题的重要性，委员会选择在本章中直接详细说明其论点、建议和实施这些建议所需的具体行动。此外，关于美国应如何调整其 TEVV 政策以保持对人工智能系统的信心的进一步细节可以在第 7 章及其相关的行动蓝图中找到，而关于国防部组织结构的相关变化的建议可以在第 3 章中找到。

第四章 – 尾注

[1] 国际人道主义法也被称为武装冲突法和战争法。

[2] 保罗 • 沙雷（Paul Scharre），《*无人之军：自主武器与战争的未来*》，诺顿出版公司（W.W. Norton & Co.），39（2018 年 4 月 24 日）。

[3] 《特定常规武器公约》中致命自主武器系统的背景，联合国（最后一次查阅日期：2021 年 1 月 11 日），https://www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument。

[4] 差别，红十字国际委员会（上次访问时间：2021 年 1 月 15 日），https://casebook.icrc.org/glossary/distinction。

[5] 在美国军事行动中，在减少目标误认方面还有改进的余地。例如，在阿富汗战争中，一项研究表明，美军部队造成的所有平民伤亡事件中，约有一半是目标识别错误造成的。使用人工智能系统做出更准确的目标选择决定，也许是人工智能的正确使用可以使战争更人性化的主要方式。拉里 • 刘易斯（Larry Lewis），*重新定义人类控制：*《自主控制战场的教训》，CNA 第 4 期（2018 年 3 月），https://www.cna.org/cna_files/pdf/DRM- 2017-U-016281-Final.pdf。

[6] 均衡，红十字国际委员会（上次访问时间：2021 年 1 月 15 日），https://casebook.icrc.org/glossary/proportionality。

[7] 参见保罗 • 沙雷（Paul Scharre），《*无人之军：自主武器与战争的未来*》诺顿出版公司（W.W. Norton & Co.）。

[8] 对于一个正确设计和测试的自主系统，正确执行指挥官的意图，指挥官显然要对该系统的行动负责。各国有责任妥善设计、测试和使用此类系统，并制定严格的程序，确保任何武器使用都符合国际人道主义法，包括确保个人问责制。

[9] 委员会认为，国防部现有“适当的人的判断”的表述，在下面的判断中讨论，抓住了这种必要的变化，并确保任何使用致命武力的决定开始于人的判断，并在人的判断的控制之下，而且人最终将对任何使用武力的决定负责任。

[10] 新闻稿，美国国防部，*美国国防部采用人工智能的伦理原则*（2020 年 2 月 24 日），https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts- ethical-principles-for-artificial-intelligence/。

[11] 国防部指令第 5000.01 号要求国防部装备的任何武器都要经过法律审查，以确保符合武装冲突法（LOAC），遵守 1949 年 8 月 12 日《日内瓦公约附加议定书》第 36 条规定之要求。国防部指令第 3000.09 号和国防部人工智能道德原则以这一基准为基础。请参阅*国防部指令第 5000.01 号：*《国防采办系统》，美国国防部第 9 期（2020 年 9 月 9 日），https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/500001p.pdf?ver=2020 - 09 - 09 -160307-310；*1949 年 8 月 12 日日内瓦公约关于保护国际性武装冲突受难者的附加议定书（第一议定书），1977 年 6 月 8 日*，红十字国际委员会（最后一次访问：2021 年 1 月 5 日)，https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/ WebART/470-750045。

[12] *国防部指令第 2311.01 号：*国防部战争法计划，美国国防部，第 11 页（2020 年 7 月 2 日），https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/ dodd/231101p.pdf?ver=2020 - 07- 02-143157-007。

[13] *国防部指令第 3000.09 号：武器系统的自主性*，美国国防部，2（2012 年 11 月 21 日，公司变更 1，2017 年 5 月 8 日），https://www.esd.whs.mil/portals/54/ documents/dd/issuances/dodd/300009p.pdf.国防部指令第 DoDD 3000.09 号规定的武器审查程序是专门为确保任何美国自主武器系统符合国际人道主义法原则，如歧视和相称性，同时也保持适当的人类判断水平并确保问责制。

[14] *国防部指令第 5025.01 号：国防部发行计划*22，（2016 年 8 月 1 日，公司变更 3，2019 年 5 月 22 日），https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/502501p. pdf?ver=2020 - 05 -20 - 081854- 657。

[15] 见本报告的附录，其中载有人工智能国家安全委员会的《负责任地开发和应用人工智能的主要考虑因素》的节选版本。关于委员会为提高人工智能系统的测试、评估、核查和验证能力而提出的未来研发建议的更多细节，参见《负责任的人工智能开发和应用的关键考虑》中“系统性能”一节：*扩展版本*，人工智能国家安全委员会（2021）（已提交委员会存档）。

[16] 《国防部战争法手册》是负责实施战争法和执行军事行动的所有国防部人员的详细资源。参见《美国国防部战争法手册》（2016 年 12 月）https://dod.defense.gov/Portals/1/Documents/ pubs/DoD%20Law%20of%20War%20Manual%20 -%20June%202015%20Updated%20Dec%202016. pdf?ver=2016 -12-13-172036 -190。

[17] 新闻稿，美国国防部，*美国国防部采用人工智能的伦理原则*（2020 年 2 月 24 日），https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts- ethical-principles-for-artificial-intelligence/。

[18] 大卫 • 阿克斯（David Axe），不要惊慌，但俄罗斯正在训练其机器人坦克听懂人类的语言，福布斯（2020 年 6 月 30 日），https://www.forbes.com/sites/davidaxe/2020/06/30/dont-panic-but-russia-is- training-its-robot-tanks-to-understand-human-speech/?sh=7373377914f2。

[19] 美国对禁止集束弹药和核武器的条约表示了类似的关切。参见问答：《集束弹药公约》，人权观察（2010 年 11 月 6 日），https://www.hrw.org/news/2010/11/06/qa-convention-cluster-munitions#；希瑟 • 威廉姆斯（Heather Williams），《禁止核武器条约对美国盟友的意义》，岩石上的战争（2020 年 11 月 5 日），https://warontherocks.com/2020/11/what-the-nuclear-ban-treaty-means-for-americas-allies/。截至 2021 年 3 月，没有一个与美国有共同防御协议的盟国表示支持禁止致命性武器系统的条约。

[20] 委员会认识到，人工智能应协助核指挥和控制装置的某些方面，如早期预警、早期发射探测和多传感器融合，以验证单一传感器的探测，并有可能消除虚假探测。

[21] 《2020 年核事务手册》，国防部负责核事务的副助理部长办公室，18，（2020 年），https://fas.org/man/eprint/nmhb2020.pdf。

[22] 各国可能还有其他原因将核武器发射权下放给自主系统，特别是如果领导层对机器执行发射命令的信任程度超过人类。一项政治协议不太可能解决这些问题，尽管提供该协议将突出其他国家如何从事不负责任和危险的行为。

[23] 迈克尔 • 派克（Michael Peck），俄罗斯的 "死神之手 "原子弹又出现了，国家利益（2018 年 12 月 12 日），https://nationalinterest.org/blog/buzz/russias-dead-hand-nuclear-doomsday-weapon- back-38492。

[24] 美国国务院新闻稿，副国务卿沙利文（Sullivan）与俄罗斯副外长里亚布科夫（Sergey Ryabkov）（2019 年 7 月 17 日）参加战略安全对话，https://2017-2021.state.gov/deputy-secretary-sullivans-participation-in-strategic-security- dialogue-with-russian-deputy-foreign-minister-sergey-ryabkov/index.html；新闻稿，美国国务院、美国和俄罗斯举行空间安全交流会（2020 年 7 月 28 日）， https://2017-2021.state.gov/the-united-states-and-russia-hold-space-security-exchange/index.html。

[25] 在过去的一年里，中国专家积极参与了与美国专家就军事人工智能系统的安全问题第 2 阶段进行的几次对话，这可能表明中国希望就这些问题进行正式的政府间沟通。.

[26] 请参见 迈克尔 C. • 霍罗威茨（Michael C. Horowitz）和保罗 • 沙雷（Paul Scharre），*人工智能和国际稳定：风险和建立信任措施*，新美国安全中心（2021 年 1 月），https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/AI-and-International-Stability-Risks-and- Confidence-Building-Measures.pdf?mtime=20210112103229&focal=none。

[27] 致命性自主武器系统领域新兴技术政府专家组 2019 年会议最后报告，《禁止或限制使用某些可被认为具有过分伤害力或滥杀滥伤作用的常规武器公约》缔约国政府专家组，CCW/MSP/2019/CRP.2/Rev.1，（2019 年 11 月 13 日至 15 日），https://undocs.org/CCW/MSP/2019/9.

[28] 见本报告的附录，其中载有人工智能国家安全委员会的《负责任地开发和应用人工智能的主要考虑因素》的节选版本。关于委员会对国际协作与合作的未来行动建议的更多细节，参见《负责任的人工智能开发和应用的关键考虑》中“系统性能”一节：*扩展版本*，人工智能国家安全委员会（2021）（已提交委员会存档）。

[29] 关于利用出口管制防止人工智能算法转让的困难，参见本报告第 14 章的补充信息。

[30] 按照这些思路，美国在 20 世纪 70 年代与苏联分享了防止未经授权核武器武装的“允许行动链接”（PAL）技术。目前还不清楚是否有一种与人工智能的允许行动链接相当的技术，这种技术可以减少人工智能系统未经授权或意外升级的风险，同时又能显著提高该系统的军事性能。如果开发出同等的技术，就必须逐一考虑合作。

第五章 人工智能和国家情报部门的未来

2025：人工智能支持的情报和预测性分析

与其他国家安全任务部门相比，情报部门将从快速应用人工智能技术中获得更大收益。随着所有可能的平台（包括机器和人）都对全球信息网络做出贡献，以及随着传感器数量的指数式增长，数据的数量、速度和种类都有可能使情报分析不堪重负，确定信息的真实性和价值将变得更加困难。分析师们将面临严峻挑战：为信息转化为可操作情报提供至关关键的背景。

人工智能将帮助情报专业人员进行“大海捞针”的工作，连接看似散乱且不相关的点，并通过辨别趋势和发现以前隐藏或掩盖的迹象和警告，破坏危险的阴谋。人工智能将改善情报周期的每一个阶段——从任务分配到情报收集、处理、利用、分析和传播。通过筛选大量数据，人工智能算法可以寻找模式、检测威胁、识别相关性并进行预测。人工智能工具可以使卫星图像、通信信号、经济指标、社交媒体数据和其他大量的信息来源变得更加易于理解。人工智能可以发现开源数据和其他情报来源之间的相关性，帮助情报部门以更加精确、高效和有效的方式进行目标定位和情报收集活动。适用于情报任务的当前和新兴人工智能技术包括：用于图像分析的计算机视觉、生物识别技术（如面部、声音和步姿识别）、自然语言处理、以及大型数据库的算法搜索和查询功能等。最重要的是，通过对不同的数据流进行数据融合，人工智能能够创建一张拼图[1]。

在军事场景中，例如对抗技术先进的对手、流氓国家或恐怖组织，人工智能情报、监视和侦察平台以及人工智能指示和警告系统，对于本报告第 3 章中讨论的先进作战能力至关重要。通过自动化，人工智能系统将以近乎实时的方式对平台、传感器和间谍的任务分配和收集进行优化，以便应对动态情报需求或环境变化。在战术边缘，“智能”传感器能够预处理原始情报，明确传输和存储数据的优先级，这在恶劣环境或低带宽环境下将起到明显作用。一旦收集完毕，智能处理系统可以对信息进行分流，识别趋势和模式，总结关键影响，并准备最高优先级的信息供人类审查（或根据分析师定义的条件，标记特别感兴趣的项目）。人工智能情报系统包括先进的指示和警告系统，它可以使作战人员更早地预测和了解新出现的威胁并积极主动地塑造环境，以及接近战术边缘、能够识别对手拒止和欺骗手段的系统等。当人工智能系统与人类的判断相结合时，这些能力将增强全域意识，产生更加紧密明智的决策周期，为不同的行动路线提供建议，并支持对对手的行动进行快速反击。

随着新技术的迅速传播，情报部门进行适应性变革的要求变得十分迫切。今天，美国情报部门昔日完美无缺的情报技巧已经在世界各地被广泛使用[2]。对手迅速应用人工智能工具的能力意味着我们可能更容易受到欺骗、信息战、情报获取来源和方法暴露、网络战和反间谍活动的影响。在人工智能应用方面，国家情报部门是联邦政府内的先行者，它很早就搭建了支持人工智能应用的一些基本的基础设施，例如：2013 年[3]，签订涵盖国家情报部门的商业云服务承包协议；2019 年，发布《使用机器增强智能的策略计划》，为更广泛的应用人工智能提供了方向和框架。此外，部分情报机构在应用人工智能方面取得了长足的进步，领先于其他政府机构。然而，在权力、政策、预算、数据共享和技术标准方面的重大障碍使情报部门无法充分发挥潜力。因此，如果不对安全审查程序进行实质性改革，本章中所给出的建议将不会达到预期效果。

雄心勃勃的时间表：到 2025 年，国家情报部门实现人工智能战备状态

为了在各个机构已经取得的成果基础上实现“百尺竿头更进一步”，国家情报部门应该制定雄心勃勃的目标，在情报事业每一个可能的方面应用和整合人工智能能力，以此作为未来情报更加宏大愿景的一部分。

到 2025 年，实现情报部门的人工智能战备状态：

情报专业人员具备基本的数字素养，能够访问在情报周期每个阶段中无处不在的人工智能集成所需的数字基础设施和软件。

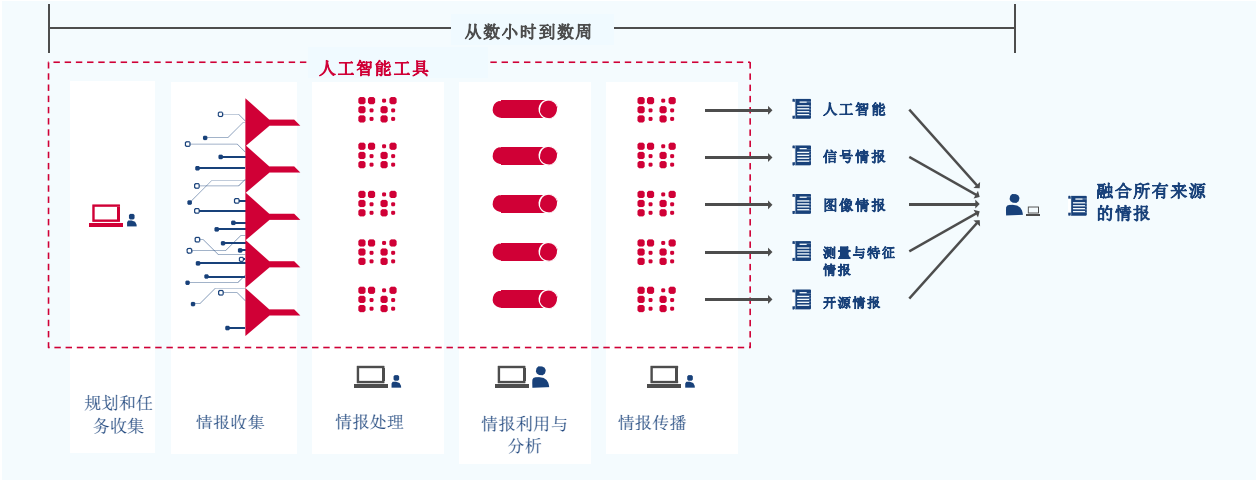
情报部门应该立即采取行动，优先推动在情报周期的每个阶段最大限度实现自动化，并在人类分析师审查之前，通过人工智能分析系统处理所有可用的数据和信息。情报产品应以机器的速度进行传播，这意味着它们必须是机器可读的格式，而且整个情报部门的系统必须能够在没有人工干预的情况下接收和使用情报产品。通过这种方式对人工智能系统进行优化，需要以完全不同的方法创建和审查成品情报产品。情报部门应该要求所有情报产品必须包括人类可读的版本以及同等重要的自动机器可读版本，后者可以被整个情报部门的其他分析系统所接收。所有未来的情报系统应该为面向人工智能的数据收集和处理进行优化。

“情报部门应该要求所有情报产品必须包括人类可读的版本以及同等重要的自动机器可读版本，后者可以被整个情报部门的其他分析系统所接收。”

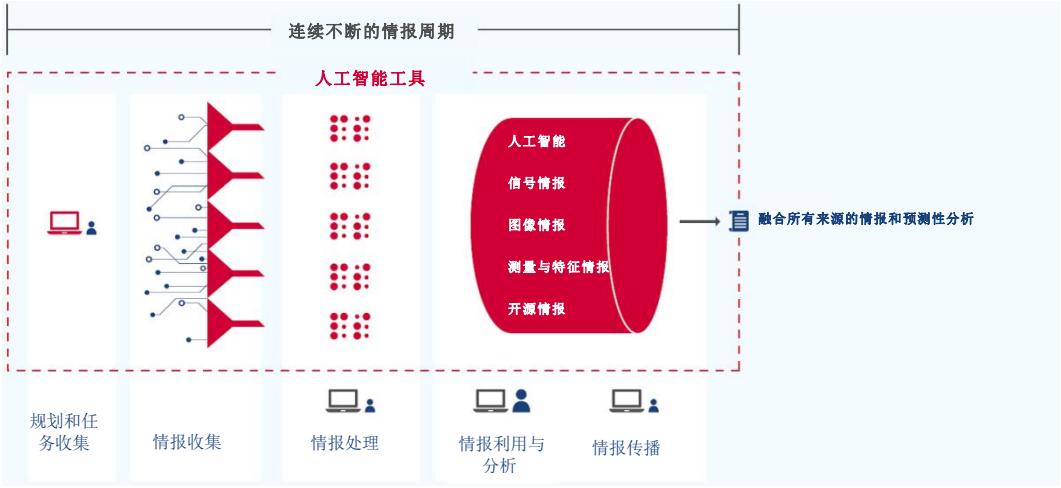
一旦情报部门在各个领域内实现流程自动化，它应该将这些单独的流程融合到一个连续的“管道”中，通过不断学习的分析引擎的联合架构处理所有来源的情报分析。这种变革使人机编组产生的洞见超越目前人类独立认知的极限。此类系统将使情报人员能够更加清晰洞见当前的事态发展，对新兴威胁进行更加准确可靠的预测分析。随着分析师对人工智能系统的信任度越来越高，以人类-机器为主导的分析比例将更多地向机器分析倾斜。

当前

优化：在当前的情报领域内，实现人工智能自动化



转变：人工智能支持所有来源的情报和预测性分析



实现“到 2025 年，情报部门的人工智能战备状态”的目标，需要采取下列行动：

建议

对情报部门内负责科技情报事务的领导进行赋权。国家情报总监应该任命国家情报总监办公室内的科技总监为情报部门的首席技术官，并授权其推动情报部门采用人工智能应用程序来满足作战情报需求。为此，首席技术官应该监督“使用机器增强智能”战略的实施情况；建立和执行必要的共同技术标准和政策，以便迅速和负责任地在整个情报部门推动使用人工智能应用程序；领导采办改革，从而确保情报部门能够迅速为情报专业人员采购和部署系统。首席技术官应被授予额外的权力来制定与情报研究和工程、技术开发、技术转移、原型样机、实验和开发测试活动相关的政策，并对上述活动进行监督。

建议

改变风险管理实践，加速推进新技术采用。情报部门需要平衡新技术上线和快速更新带来的技术风险，与因为没有跟上发展步伐导致的重大作战风险（国防部也面临着类似的窘境）。与此同时，应尽可能地实现软件定期升级实现自动化。为了在各机构之间更加便捷地共享软件工具，信息技术系统的相互认证应该成为标准[4]。

“情报部门需要平衡新技术上线和快速更新带来的技术风险，与因为没有跟上发展步伐导致的重大作战风险……”

为了协调这些变化，国家情报总监办公室应该建立一个专注于技术现代化的高级风险管理委员会[5]。委员会的任务是权衡采用新技术的风险和不采用新技术的机会成本。委员会的目标是确保分析师能够获得工作所需的工具。

情报部门将需要国会情报委员会的支持，例如在更灵活的软件开发框架内灵活使用资金。为了支持提高灵活性的论点，情报部门应该开发数据驱动的方式来向外界清晰展示运营收益，以及对不作为的风险进行可信的评估。

建议

改善情报部门与国防部之间的协调和互操作性。情报部门必须积极寻求与国防部之间的自动化互操作性，以便在机器速度下进行情报操作[6]。为了实现这一目标，安全管理人员和网络管理员必须对快速安全的数据交换建立更大的信心。国家情报总监办公室、负责情报和安全的国防部副部长以及联合人工智能中心应该在与情报有关的人工智能项目上进行更多协调，尽可能减少重复性工作，同时最大限度地利用人工智能能力开发、测试和评估、部署和国际参与的通用方法以及相关政策和权力。三者应该合作创建可互操作和可共享的资源和工具，即本报告第 2 章所述的人工智能研发生态系统中所设想的宏大愿景，并应在可行情况下，建立共享所有人工智能能力的文化[7]。

建议

利用人工智能对开源信息和公开信息进行分析[8]。情报部门应该制定一种协调联合的方法，将人工智能用于开源情报，并在每个情报领域内，尽可能地将开源分析纳入现有的情报流程[9]。

建议

优先考虑并加速科技情报收集工作，从而更好地了解对手的能力和意图。情报收集工作要求情报部门大幅提高分析人员的技术精度、能力和潜力。相应的举措必须包括积极努力地培训、招聘和保留具有必要技能的分析师——他们必须明确指导收集要求，并提供及时、准确的评估。为了更好协调科技情报的收集工作，包括收集竞争对手之间进行的科技合作，国家情报总监应在国家情报委员会内任命一位新兴技术收集主管[10]。

建议

为了招募更多的科技专家进入情报部门，积极推行绝密级及以上级别的安全审查改革，并在情报部门工作人员内部实行相互安全审查。国家情报总监办公室应该制定并实施基于人工智能数据和科学的安全审查裁决办法，缩短调查时间[11]。

建议

推进并继续开发专门的信息技术环境，它能够融合来自不同领域和来源的情报。人工智能支持的技术架构可以协助自主整合不同领域的情报，而目前往往需要人工干预来分享原始数据或分析结果[12]。这种做法有助于情报部门融合产生于不同信息流的洞见，创建一张拼图。举例来说，信号情报往往依赖人类情报或地理空间情报。同样，人类情报的价值几乎总是可以通过对其上面的信号情报或开源信息进行分层来得到加强。

建议

接受经过融合的预测性分析作为新标准。成功融合所有来源/领域的情报将使准确的预测性分析成为可能，而目前这还是一个无法达到的目标。政府对新冠肺炎的反应让我们看到了融合数据集并为这种分析提供信息的可能。举例来说，美国北方司令部（与联合人工智能中心和国民警卫局合作）根据从几十个不同的数据集中建立的预测模型，识别新冠肺炎的热点并协调对重要物资的需求[13]。

建议

开发以人为本的创新方法来进行人机编组。此处设想的通过机-机自主整合进行的数据融合需要全新的人机编组理念，从而实现双方各自优势的优化[14]。情报部门需要通过新的方法放大和扩展人类的认知，以便有效处理由所有来源的情报分析引擎导致的信息规模和复杂性。在开发此类系统时，情报部门必须充分了解并审慎决定人类或机器应该独立行动，以及人类和机器应该编组合作的时间和条件。.

“此处设想的通过机-机自主整合进行的数据融合需要全新的人机编组理念，从而实现双方各自优势的优化。”

第五章 – 尾注

[1] 关于整个情报周期中的人工智能使用案例的其他信息，参见《维持情报优势中的“应用”：*通过创新来重新想象和重塑智能》的讨论*，战略与国际研究中心技术和情报工作组，8-22（2021 年 1 月 13 日），https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/210113_Intelligence_Edge.pdf。

[2] *使用机器增强智能的策略计划：*《使用机器增强情报的战略》，国家情报局长办公室（2019 年），https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf（国家情报局首席副局长休 • 戈登（Honorable Sue Gordon）阁下的前言）。

[3] Frank Konkel，中情局与亚马逊交易的细节，大西洋（2014 年 7 月 14 日），https://www.theatlantic.com/technology/archive/2014/07/the-details-about-the-cias-deal-with- amazon/374 632/.

[4] 在采用新的软件系统时，情报机构（IC）遵循国家标准与技术研究所（NIST）制定的风险管理框架。虽然，整体上这是一个有用的框架，但它也会造成延误，或使集成电路跟不上商业上的尖端人工智能工具。有关更多信息，请参阅 FISMA 实施项目，NIST（2020 年 12 月 3 日），https://csrc. nist.gov/projects/risk-management/rmf-overview。

[5] 高级风险管理委员会将帮助美国情报机构执行拟议的新兴技术三主席委员会的指导意见，其功能类似于本委员会建议负责研究和工程的国防部副部长作为联合需求监督委员会的共同主席的作用。

[6] 有关更多信息，请参见肯特 • 林尼布尔（Kent Linnebur）等人的《下一步后的情报：未来的情报机构工作场所，米特尔技术和国家安全中心（2020 年 11 月 1 日），https://www.mitre.org/sites/default/files/publications/pr-20 -1891-intelligence-after-next-the-future-of-the-ic-workplace.pdf。

[7] 这些努力应利用联合人工智能中心的联合共同基金（JCF）。

[8] 出版 L. 116-260，《综合拨款法》（2021 年），W 分部，第 326 节（"情报界的开源情报战略和计划"），第 623 节（“关于开源情报的独立研究”），和第 624 节（“关于开源企业的调查”）为国际刑事法院重新想象开源情报的作用提供了一个起点。

[9] 必须指出，开放源码情报（OSINT）不限于传统媒体来源（报纸、广播等）和社交媒体。OSINT 还包括公开资料，如政府公共数据源（官方报告、预算文件、听证会证词等）、专业和学术出版物、商业数据源（行业报告、财务报表、商业图像等）等等。

[10] 更多信息，请参见《保持情报优势：*通过创新重塑情报》中关于 "提升技术情报 "的讨论，战略与国际研究中心技术和情报工作组，12（2021 年 1 月 13 日）*，https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/210113_Intelligence_Edge.pdf。

[11] 关于需要对安全审查裁决的行为方法进行学术和科学审查的更多信息，见大卫 • 勒基（David Luckey）等人，评估内部威胁的持续评估方法：如何改善美国各部门和机构的安全态势？兰德公司，第 28-34 页（2019 年），https://www.rand.org/pubs/research_reports/ RR2684.html.

[12] 本报告第 2 章更为详细地介绍了这种环境的技术问题。

[13] 空军上将特伦斯 • 奥肖内西（Terrence J. O’Shaughnessy）、美国北方司令部司令、陆军中将劳拉 • 理查森

（Laura J. Richardson）、美国北方陆军司令，文字记录：美国北方司令部和阿诺思指挥官讨论正在进行的COVID-19 努力，美国国防部（2020 年 4 月 21 日），https://www.defense.gov/Newsroom/ Transcripts/ Transcript/Article/2160070/us-northcom- and-arnorth-commanders-discuss-ongoing-covid-19 -efforts/。

[14] 参见肯尼斯 • 福特（Kenneth M. Ford）等人，*认知矫正*：走向以人为中心的人工智能，《人工智能》杂志第7期（2015年冬季），https://doi.org/10.1609/aimag.v36i4.2629; John Laird等人，美国国防部人机合作研讨会未来方向（2019年7月16-17日），https://basicresearch. defense.gov/Portals/61/Future%20Directions%20in%20Human%20Machine%20Teaming%20 Workshop%20report%20%20%28for%20public%20release%29.pdf；Gagan Bansal等人，最精确的人工智能是最好的队友吗？团队合作优化人工智能，AAAI 2021（2021年2月），https://www. microsoft.com/en-us/research/publication/is-the-most-accurate-ai-the-best-teammate-optimizing-aifor-teamwork/.

第六章 政府中的技术人才

提高政府中技术人才的素质

在这场人工智能竞赛中，胜利不会属于拥有最好技术的一方，而是属于拥有最出色、最多样化和最精明能干的技术人才的一方。国防部和情报部门都面临着惊人的人才短缺。目前，这一问题已经成为“到 2025 年，美国实现人工智能战备状态”的最大障碍。国家安全机构需要更多的数字人才，否则它们在采购、部署和使用人工智能和相关技术方面仍然无法做好准备。数字人才是实现政府现代化最重要的要求，但是截至目前，很少有部门对打造一支数字化的人才队伍进行充分投资[1]。

“国防部和情报部门都面临着惊人的人才短缺。”

为了扩大数字化和人工智能人才队伍，政府需要：

- 通过旨在容纳高技能专家的人才管理系统，有效**组织**政府内部的技术专家；
- **招募**已经拥有政府所需技能的人才，如行业专家、学者和大学应届毕业生；
- 通过培训和教育当前和未来的政府雇员，**打造**自身的数字人才队伍；
- 更有效地实施数字人才**雇佣**政策，确保他们在进入政府之后，能够从事富有意义的工作。

“数字人才是实现政府现代化最重要的要求……”

当前模式

负责创建人工智能解决方案的政府组织正在努力打造自身的数字人才队伍。不过，政府组织内部的障碍使得它们无法招募并留住人工智能从业者和更广泛领域内的数字人才。在薪资水平上，政府无法与私营部门展开竞争；在招聘流程上，政府不得不实施冗长繁琐的招聘过程；而且，所有改革措施都受困于缓慢的安全审查流程。

对于这种现状，我们不能无动于衷，或者将其看作不可避免的未来。政府可以与私人部门开展人才竞争。在薪资水平上，政府可能无法匹配私营部门的薪水，但是它能够提供应对国家安全挑战以及为社会做出重大贡献的机会。事实上，阻碍数字人才招聘的最大障碍并不是薪酬问题。政府部门普遍流行的看法，甚至往往是政府部门的现状，造成内部数字人才很难利用现代化的计算工具，在快速变化的数字技术领域前沿开展有意义的工作[2]。

“对于这种现状，我们不能无动于衷，或者将其看作不可避免的未来。政府可以与私人部门开展人才竞争。”

有观点认为，政府应该专注于项目管理和数据收集与管理，并将所有开发工作进行外包。我们从部分领导人那里听到了这种观点，他们认为政府雇用或培训自己的人工智能专家的做法不切实际。有趣的是，我们从未从产业界听到这种说法。人工智能国家安全委员会对这种看法持反对意见。

> “任何不培养和发展政府技术人才队伍的战略都是短视的。”

任何不培养和发展政府技术人才队伍的战略都是短视的。仅仅依靠承包商提供数字专业知识的政府机构将无法很好地理解基础技术，从而无法独立于承包商做出成功的采购决策[3]。这种情况会造成国家安全风险。尽管承包商应该继续发挥关键作用，它们被激励，在某些程度上是被要求履行合同条款，而不是寻求整体系统的改善，或对不够成熟的要求或无效的战略提出反对意见。因此，依靠承包商的政府机构内的数字专家在关键决策中，甚至在与他们的专业知识领域相关的决策过程中，不得不退居次席。美国政府永远都会有自己的承包商。但是，美国政府可以而且更应该培养和发展自己的数字人才队伍。

组织

数字人才队伍的组织方式与其专业知识水平一样重要。为了能够按照国家安全事业所需的规模形成和管理一支熟练的数字人才队伍，政府需要建立相应的人才管理框架。

建议

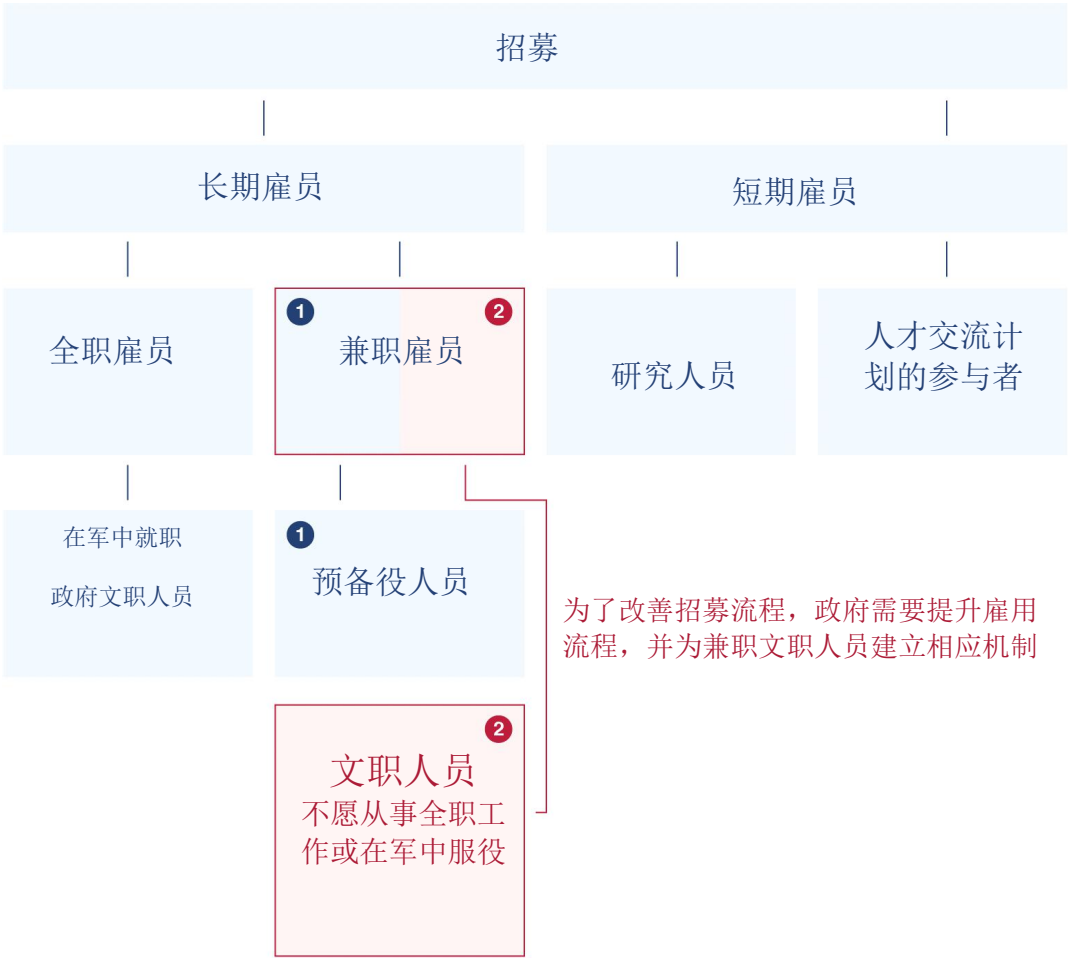
政府部门和特定机构应该组建数字部队。我们建议政府部门和特定机构应该组建数字部队，开展人员招募、培训和教育；将人员放在数字岗位上或从数字岗位上移除；管理人才的数字职业生涯；设置数字人才资质标准。政府部门和特定机构将为数字部队中的成员设置岗位，并向其提供上岗指导。

数字部队模式受到了陆军医疗部队的启发，后者汇聚了一批不适合陆军传统人才管理框架、但是又具备专业医疗保健技能的专家。与医疗部队一样，政府机构的数字部队应该具有专业的人才政策、晋升准则、培训资源和能够证明部队成员在全新数字领域内的熟练程度的人才认证体系。

招募

为了填补数字部队的成员空缺并提升其更广泛的数字人才队伍的素质，政府需要改进招募和雇用流程，加速推进安全审查，利用《政府间人员法》等临时雇用政策工具，建立兼职文职机制[4]。许多人工智能和其他数字从业者有兴趣以全职雇员或兼职雇员身份与政府进行合作。在希望获得全职工作的候选人中，部分人士寻求在整个职业生涯从事文职工作或在军中服役。而其他人士则不太愿意做出长期承诺，相反他们更希望成为短期的全职雇员、研究人员、人才交流计划的参与者或军队的预备役人员。第三类人愿意与政府合作或为政府兼职工作，但是他们无意成为全职文职雇员或在军中服役。

招聘模式上的差距



建议

建立面向文职人员的国家后备数字部队。政府应该通过建立相应的雇用机制，挖掘愿意为公共服务贡献部分时间的技术专家。虽然兼职员工无法替代全职员工，但是他们可以协助改善人工智能教育，进行数据分流和获取，帮助指导项目和构建数字解决方案，在公共和私营部门之间建立桥梁，并承担其他重要任务。

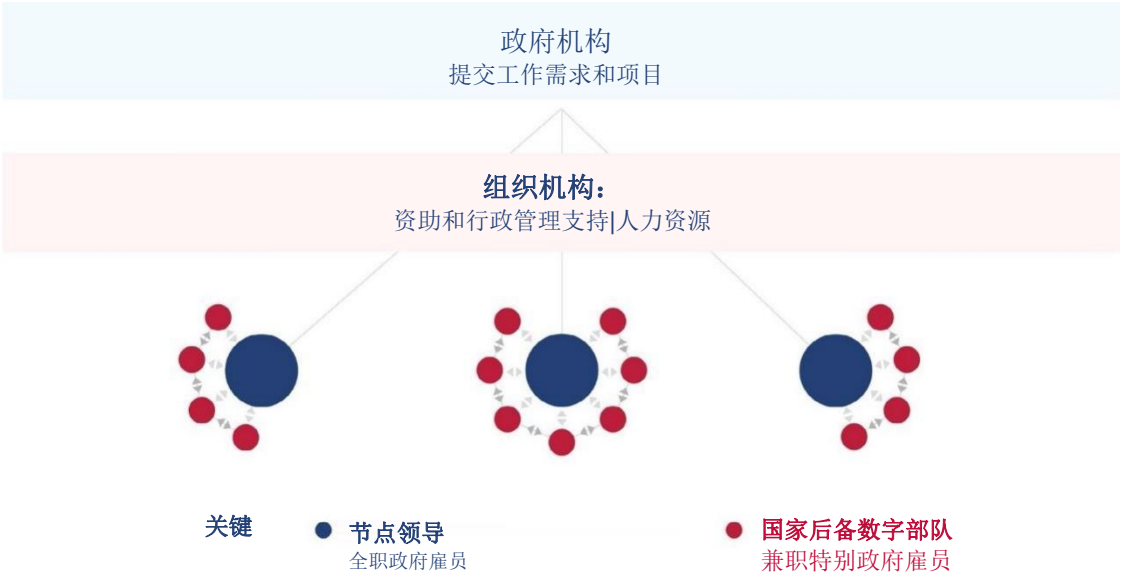
为了消除招聘缺口，政府应该效仿军队预备役的责任和激励结构，建立面向文职人员的国家后备数字部队。部队成员将成为某个机构数字部队的政府特别文职雇员，每年以顾问、教师或开发人员的身份，为政府服务至少 38 天。

建议

简化雇用流程，扩大数字人才招聘渠道。美国政府雇用制度存在的问题众所周知：进展缓慢，努力招聘技术专家，但是年轻或没有学位的专家，特别是当他们申请专业水平和工资水平相匹配的岗位时，却很难受聘。这些挑战并不是由缺乏招聘权限或招聘过程过于缓慢造成的。委员会无法识别数字人才招聘权限上存在的空白。

为了消除招聘瓶颈，政府需要针对大学和政府服务部门，扩大科学、技术、工程和数学以及人工智能人才的招聘渠道，简化雇用流程，并为数字部队或政府机构创建服务于特定机构和军种的数字人才招聘办公室。招聘办公室将监测数字部队、政府机构或军种对特定类型的数字人才的需求，并根据相应授权，通过参加会议和招聘会、在大学校园进行招聘、举办有奖竞猜、提供奖学金、招聘奖金和推荐奖金等方式招聘技术专家。

后备数字军团.



> 委员会无法识别数字人才招聘权限上存在的空白。

常设数字部队将监督整个政府的招聘进展，并提出建议，以便扩大和改善数字人才的雇用流程和招聘渠道。他们也应该能够尝试执行新的权限。

打造

政府将无法通过招聘满足技术人才缺口。人工智能和数字人才在全美境内十分稀缺。2020 年，全美计算科学岗位空缺超过 43 万个，但是每年新毕业的计算机科技人才只有 71,000 名[5]。政府应该做出全新承诺，通过采取重大举措，从根本上打造数字人才队伍。

> “美国需要建立全新的服务型学院，培养具备数字技能的未来文职人员，推动政府的现代化进程。”

建议

成立美国数字服务学院。美国需要建立全新的服务型学院，培养具备数字技能的未来文职人员，推动政府的现代化进程。 美国数字学院将是一所政府备案并且授予学位的大学，接受政府部门资助和私人捐助，由联邦政府内部专门成立的独立机构进行管理，满足政府对数字专家的需求（具体需求由一个机构间委员会确定），并由私营部门和学界技术领袖组成的联邦咨询委员会进行协助。美国数字服务学院应仿照五大军事院校的模式，为所有联邦政府部门和机构培养受过训练和教育的政府文职人员。

国数字服务学院执行计划建议.

第一阶段

（第 1-2 年）

- 建校选址，并为未来的规模扩大预留空间
- 在人事管理办公室的领导下，一个机构间特别工作组确定政府目前的数字人才队伍和理想的数字人才队伍之间的差距
- 将美国数字服务学院的管理层设立成为联邦政府下属的一家全新的行政机构，能够获得单独的拨款，负责分阶段实施计划和机构的管理
- 招聘终身制教员
- 主要面向私营部门的技术公司招聘兼职教师
- 授权美国数字学院接受个人和企业的外部资助和礼物，用于启动、维护和基础设施费用
- 拨款 4 千万美元，用于支付管理成本。
- 满足教育部以及美国数字教育学院所在州规定的必要要求，获得授予学位的批准。
- 向工程技术评审委员会下属的大学计算科学鉴定委员会申请学位课程认证
- 向教育部和高等教育认证委员会批准的区域认证组织申请认证，获得“候选”资格
- 构建最初的实体性基础设施
- 额外进行拨款，用于选择和购买校园土地及建造基础设施

第二阶段

（第 3-5 年）

- 在第三年开始时进行授课，第一届学生为 500 人
- 证明符合区域认证组织的所有要求和标准，以便获得成员资格。

第三阶段

（第 6-7 年）

- 第一届学生毕业
- 通过认证评估项目持续进行改进
- 评估并酌情扩大班级规模

“数字人才应该能够合理预期：在职业生涯内，他们将在联邦政府内，专注于专业领域，从事有意义的工作。”

雇用

数字人才应该能够合理预期：在职业生涯内，他们将在联邦政府内，专注于专业领域，从事有意义的工作。如果没有这样的职业预期，他们不可能加入美国政府的数字人才队伍；如果技术经验无法与职业预期相匹配，他们也无法长期从事相关工作。调整针对数字人才队伍的职业预期和经验背景，需要以下三点改变：

- 技术人员有机会在整个职业生涯内专注于所热爱的领域；
- 博学多识的领导者，部分领导者本身就精通数字技术；
- 获得工具、访问数据集和使用基础设施。

事实上，真正的变化远远比上面描述的更具战术性，而且影响也不小。战略举措在战术层面上取得成功或遭遇失败。许多本来可能具有战略影响的数字举措在战术层面上陷入困境或遭遇失败，部分原因在于政府未能充分发挥技术人员的才能。

建立全新的数字职业领域。对于一个组织而言，全新的职业领域对于其必要能力的定义，甚至其身份的本质，构成了挑战。如果军事部门为软件开发人员和数据科学家创建职业领域，这将几乎不可避免地会改变作为一名士兵、水手、飞行员或海军陆战队员的含义，就像数代以前，人类将航空引入作战领域一样。政府应该在软件开发、软件工程、知识管理、数据科学和人工智能方面创建民用职业类别。军事部门应该在软件开发、数据科学和人工智能方面创建职业领域，并将其划分为管理类别和技术专家类别。数字部队在发展过程中需要更多的职业领域，而建立这些职业领域将为推动政府现代化打下坚实基础。

建议

扩大获得工具、访问数据集和使用基础设施的机会。就职于政府部门或机构的高技能技术人才常常无法使用软件工程工具。数字人才队伍需要获得与私营部门相同的、使用企业级软件的能力，包括软件工程工具、访问软件库、经过审查的开源支持、策展数据集以及大规模合作的基础设施。

所有职业领域都需要改善对开源库和工具的访问6。大多数先进的人工智能和机器学习库需要海量数据来训练模型。向人工智能从业者提供横跨物理和生物科学、经济学和行为研究的丰富数据集，将使他们能够专注于自己的专业领域，无需从晦涩的数据源查找数据。

第六章 – 尾注

[1] 在政府的一些部门，如美国数字服务、凯赛尔运行、陆军人工智能工作队、美国空军-麻省理工学院人工智能加速器、情报界的组成部分和国家实验室，都有一些优秀的小区域，但数量太少，而且在政府中的传播不够广泛。各机构对人工智能队伍的规模和类型的要求各不相同，但人工智能国家安全委员会（NSCAI）接触过的每个机构都表示需要扩大其人工智能队伍，这里的建议也广泛适用。

[2] 人工智能国家安全委员会（NSCAI）的员工与防务创新委员会和国防数字服务局的讨论（2019 年 5 月）。

[3] William A. LaPlante，拥有技术基线，国防 AT&L，18-20（2015 年 7 月-8 月），https://apps.dtic.mil/dtic/tr/fulltext/u2/1016084.pdf。

[4]有关《政府间人事法》的更多信息，请参见《政府间人事法》，OPM（最后一次查阅日期：2021 年 2 月 1 日），https://www.opm.gov/policy-data-oversight/hiring-information/ i ntergovernment-personnel-act/。.

[5] code.org（上次访问时间：2021 年 1 月 11 日），https://code.org/promote。另请参阅奥伦 • 埃奇尼（Oren Etzioni），*特朗普在人工智能方面的行政命令缺失：*美国需要一个特别的签证计划，旨在吸引更多的人工智能专家和专家，连线（2019 年 2 月 13 日），https://www.wired.com/story/what-trumps- executive-order-on-ai-is-missing/.

[6]尤其是，人工智能职业领域，TensorFlow 是世界上最流行的训练神经网络和其他机器学习（ML）算法的库之一。PyTorch 是另一个开放源代码库，它有助于将研究原型转换为可投入生产的机器学习模型。

第七章 对于人工智能系统建立合理的信心

对于采用和部署人工智能建立合理的信心

领导层

问责制和治理

稳健可靠的人工智能

人类-人工智能交互和编组

测试、评估、核查和认证（TEVV）

人工智能系统必须以合理的信心进行开发并投入使用[1]。如果人工智能系统不能按设计工作，或在可能产生重大负面影响方面无法进行预测，那么领导人将不会采用它们，作战人员将不会使用它们，国会将不会资助它们，美国人民也不会支持它们。

“如果人工智能系统……在可能产生重大负面影响方面无法进行预测，那么领导人将不会采用它们，作战人员将不会使用它们，国会将不会资助它们，美国人民也不会支持它们。”

实现可接受的人工智能性能往往与可接受一定程度的风险决策有关。没有任何一种技术能够在所有条件下都完美运作。风险计算随情况变化。在军事、情报、国土安全和执法任务中，当依赖人工智能进行判断时，所依据的变量和考虑因素通常存在较大差异。举例来说，在类似战斗等高威胁环境中，采用能够提供即时军事优势的系统在某些情况下可能比较合理，但是同时必须承认因为环境的原因，系统可能会出现失灵的现象；然而，在其他情况下，一名能够进行理性判断的指挥官可能希望在生命受到威胁时，在部署人工智能系统前得到关于系统可靠性的最高保证。

“随着政府部门和机构对机器的依赖程度越来越高，整个国家安全场景的核心指导原则是人类判断持续维持中心地位。”

随着政府部门和机构对机器的依赖程度越来越高，整个国家安全场景的核心指导原则是人类判断持续维持中心地位。负责使用人工智能的工作人员需要对风险、机会和权衡有着清晰的认知。他们需要意识到系统预期性能存在多种可能性和局限性。最终，他们需要针对“在给定的情况下，对系统多大的信心才能被称作足够的信心”这一问题，有理有据地给出答案。从系统采购、开发以及部署特定的人工智能密集型系统的合理信心阈值，到系统在战场上的表现，这些问题始终贯穿并影响人工智能系统的整个生命周期。

在使用人工智能系统这一问题上，我们无法给出绝对完美的保证，但是一些政策性因素和最佳实践能够确保我们以负责任地态度进行决策。政府机构已经广泛意识到部署人工智能系统面临的主要挑战，以及在人工智能系统工程和管理方面引入最佳实践的必要性。

委员会已经详细制定了国家安全部门以负责任的态度开发和部署人工智能的指导框架（见附录：负责任地开发和部署人工智能时的关键考量因素）。框架包含专为政策制定者和技术从业者制定的关键考量因素，涵盖人工智能的整个生命周期。与此同时，框架还包括委员会推荐的一些实践方法。随着技术的不断进步，这些实践方法应不断地进行整合与更新。目前，一些部门已经开始采取行动，整合框架中推荐的实践方法[2]，这是对委员会工作的高度认可与褒奖。

为了帮助政府机构达到负责任的人工智能的基线标准，我们在框架中重点阐述了五个问题领域的主要挑战，并给出关键性建议。

1. 稳健可靠的人工智能

目前的人工智能系统，例如用于感知和分类的系统，都出现了不同类型的失灵现象，主要表现为“误报率”和“漏报率”。当人工智能系统运行在性能边界时，它们往往表现的十分脆弱，并且我们很难预测其能力边界[3]。它们也很容易受到攻击，并在运行过程中表现出不必要的偏见。就国家安全任务而言，这些问题是十分严重的瑕疵。因此，政府机构应该：

建议

将更多联邦研发资金用于提升人工智能的安全性和稳健性。投资还应改善人工智能系统的可解释性，以便用户能够更好地了解系统是否按预期运行。

建议

咨询跨学科的专家小组，开展风险评估，改进文档编制实践，通过构建整个系统架构来限制系统失灵的后果[4]。系统架构应该能够安全检测组件性能，并且在发现异常时，能够处理错误[5]；包含能够进行自我保护（验证输入数据）和自我检查（验证向系统其他部分传递的数据）的人工智能组件；包含积极的压力测试。

“政府需要人工智能系统来增强和补充人类的理解和决策，从而使人类和人工智能的互补优势能够作为一个最佳团队得到利用。目前，实现这一点仍然面临挑战。”

2. 人类-人工智能交互和编组

政府需要人工智能系统来增强和补充人类的理解和决策，从而使人类和人工智能的互补优势能够作为一个最佳团队得到利用。目前，实现这一点仍然面临挑战。举例来说，人类容易过度信任机器，或对机器信任不足，这取决于具体的环境。目前在以下方面仍存在巨大挑战： 测试人类-人工智能编组性能，传递足够信息同时避免认知过载，使人类和机器能够理解交换控制权的具体条件，保持适当的人类参与持续了解情况并在必要时采取有意义的行动等。政府机构还需要确定与人类相比之下的机器性能标准和期望。政府应该：

建议

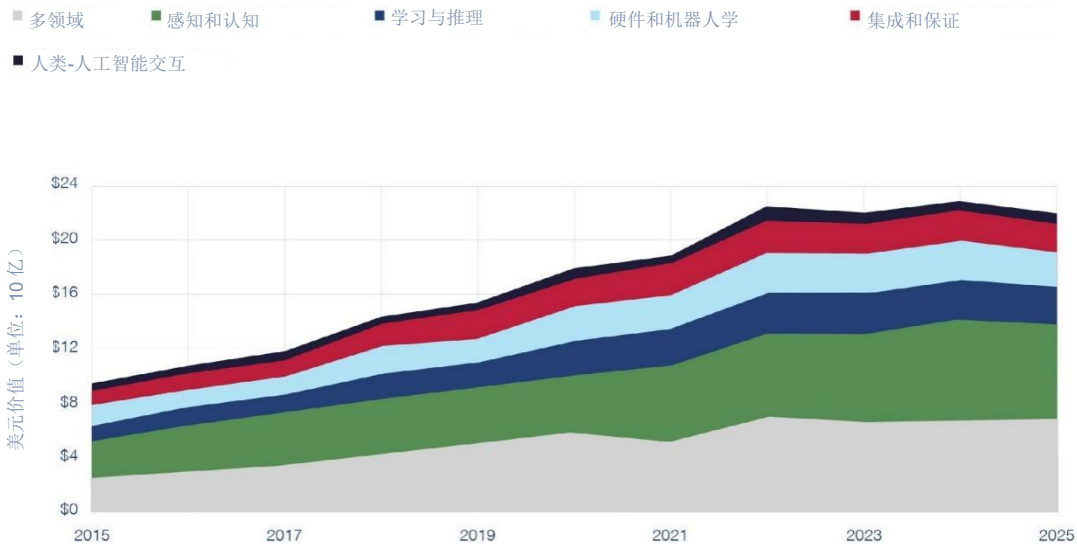
*通过国家安全研究实验室，寻求一项可持续的、跨多学科的、旨在加强人类-人工智能编组性能的举措。*这一举措倡议应该侧重于最大限度地发挥人类-人工智能互动的优势；在与人工智能系统合作时，包括通过与终端用户的持续接触和实验，更好地测试人类的表现和能力；以及帮助人工智能系统更好地理解情况的细微差别。

建议

明确人类角色和功能的相关政策，开发能够优化人机交互的设计，提供持续性的和全组织的人工智能培训。

按研究领域划分的国防部的人工智能研发、测试和评估（RDT&E）总投资，2015-2025 财年

来源：戈维尼（Govini）



国防部人工智能RDT&E总投资

这份图表显示了国防部在人工智能国家安全委员会设计划分的五大研究领域内的支出水平估值。根据图表，我们可以看到国防部在人类-人工智能交互领域内的投资远远落后于其他研究领域。

注意：表中展示的支出水平是基于对国防部 2021-2025 财年 RDT&E 预算记录分析结果得出的估值。参见《国防部人工智能 RDT&E 投资分析》，人工智能国家安全委员会（最终在委员会进行存档备案）。由于源数据中固有的质量问题，表格中的估值包含重大且难以估计的误差幅度。

3. 测试、评估、核查和认证（TEVV）

对人工智能系统建立合理的信心，必须保证它们会按照预期的方式运行，包括在与人类和其他系统进行交互时。对于传统老旧设备进行测试、评估、核查和认证无法提供充足的保证。因此，政府机构缺少共同的衡量标准来评估系统将按预期方式运行的可信度。为了最大限度减少性能问题和无法预期的结果，我们需要全新的 TEVV 流程。这是一项充满挑战的优先级任务。联邦政府需要增加研发投资来提升我们对于如何开展与人工智能和软件相关的 TEVV 工作。为了实现这一目的：

建议

随着国防部人工智能系统在数量、范围和复杂性上有所提升，国防部应该调整和制定 TEVV 政策，开发 TEVV 能力，以便满足人工智能所需的变化。这应该包括建立一个整合了持续测试的 TEVV 框架和文化；使 TEVV 工具和能力在国防部内部更加容易获得；更新或创建人工智能系统的现场、虚拟和建设性测试的范围；以及重组系统设计、开发和测试需求的基本流程[6]。

建议

国家标准与技术研究院应该提供并定期更新一整套标准、性能指标和工具，以便对人工智能模型、数据、受训环境和预测性结果建立足够信心。国家标准与技术研究院应该带领人工智能界建立上述标准、性能指标和工具，与来自产业界、学术界和政府的专家和用户密切接触，确保标准、指标和工具的有效性。

4. 领导层

负责任地开发和部署人工智能需要终端用户、高级领导人对于系统能力和局限性有着清醒的认识，从而确保人类智能系统不会滥用。它还需要专家在培训、采购、风险评估和采用不断发展的最佳实践方面提供支持。目前，在负责任的人工智能领域，只有国防部一家机构在全心全意地执行领导责任；国家安全机构中的员工通常以自愿和兼职的方式承担这些角色。如果没有全职的专业工作人员，政府机构在全面采用和实施报告中给出的建议时，无法取得成功。因此，政府应当：

建议

在对国家安全至关重要的每个政府部门或机构，以及武装部队中各自任命一名全职的、高级别的、全面主管负责任的人工智能相关事务的领导。他/她应该推动负责任的人工智能培训，提供有关负责任的人工智能政策和实践的专业知识，负责机构间协调，制定采购政策。

建议

在国家人工智能倡议办公室建立一个由多学科专家组成的常设机构。常设机构将根据实际需要，就负责任的人工智能问题向政府机构提供建议。常设机构成员应包括在人工智能和伦理学、法律、政策、经济学、认知科学和技术（包括对抗性人工智能技术）等交叉领域内拥有专业知识的人才。

5. 问责制和治理

国会和公众需要看到，政府有能力及时发现和修复系统中的关键缺陷，从而防止无意中引发的灾难；追究相关人士的责任，包括滥用人工智能系统的责任。政府机构需要具备足够的能力：在系统运行时，监测人工智能的性能（以评估它们是否按预期运行）；构建装配必要仪器设备的系统来达到上述目的[7]。对国家安全至关重要的政府部门、机构及监督实体都表示在系统可见性方面存在挑战，而供应商则呼吁明确仪器以及可审核性的要求。因此，政府机构应该：

建议

调整和扩展现有的问责政策，以便涵盖人工智能系统及其组件的全生命周期。

建议

*制定政策，允许个人对不负责任的人工智能开发提出关切，并制定全面的监督和执行措施，*包括：审计和报告要求、对最敏感或高风险的人工智能系统的审查机制，以及对受人工智能系统行动影响的各界人士的上诉和申诉程序。

对于应用和部署人工智能系统建立合理的信心：

对人工智能系统建立合理的信心需要确保人工智能应用、使用、资助和公共信任

领导层

在对国家安全至关重要的每个政府部门或机构以及武装部队中各自任命一名全职的、高级别的、全面主管负责任的人工智能相关事务的领导。

国家人工智能倡议办公室应该组建一个由多学科专家组成的常设机构，就负责任的人工智能问题向政府机构提供建议。

问责制和治理

政府机构应该调整和扩展现有的问责政策，以便涵盖人工智能系统及其组件的全生命周期；制定政策，允许个人对不负责任的人工智能开发提出关切，并制定全面的监督和执行措施。

稳健可靠的人工智能

政府机构应该咨询跨学科的专家小组，开展风险评估，改进文档编制实践，通过构建整个系统架构来限制系统失灵的后果

人类-人工智能交互和编组

国家安全研究实验室应该寻求一项可持续的、跨多学科的、旨在加强人类-人工智能编组性能的举措。政府机构应该明确人类角色和功能的相关政策，开发优化人机交互的设计，提供持续性和全组织的人工智能培训。

第七章 – 尾注

[1] “合理的信任”一词取自一个广泛使用的国际标准，它使用了一个具体的定义，即保证是“合理的信任的理由”。其指出，“利益相关者在依赖一个系统之前，需要有合理的信任理由”，“依赖的程度越大，越需要有强有力的信任理由”。ISO/IEC/IEEE 国际标准-系统和软件工程-系统和软件保证，IEEE/ISO/IEC 15026-1（2019），https://standards.ieee.org/standard/15026 -1_ Revision-2019.html.

[2] 国防部联合人工智能中心（JAIC）负责任的人工智能小组委员会和测试与评估小组委员会都进行了大量的摸底工作，以确定哪些现有做法与《关键考虑因素》中的建议相对应。国土安全部（DHS）的人工智能战略中的内容也加强了关键考虑因素的建议。参见国土安全部人工智能战略，国土安全部（2020 年 12 月），https://www.dhs.gov/publication/us-department-homeland-security-ar tificial- intelligence - strategy ? topic=intelligence - and - analysis.

[3] 像其他智能系统（包括软件和人类）一样，人工智能系统也有能力限制。但是，我们在了解人工智能系统的性能限制方面的科学性较低，包括为什么、何时以及如何失败。

[4] 这样的跨学科团队应该探索文档/标签的可能性，指定系统设计和测试的狭义任务/任务的可能性。如《负责任开发和制造人工智能的关键考虑事项附录》所述，人工智能生命周期的文件应包括用于训练和测试模型的数据信息以及用于测试模型的方法，两者都是基于使用环境。也应该包括在不同的场景或设置中使用系统时重新测试、再培训和微调的需求。

[5] 监控可以增加一层稳健性，但也必须保护自身，以防止外部间谍活动或篡改人工智能系统的新机会。

[6] 如本报告第 2 章所述，将需要对数字基础设施进行升级，以增强物理试验范围，创造能够利用数字孪生的数字测试环境。

[7] 增加新的传感器和仪器的情况下，同样可以引入新的漏洞。尤其重要的是，要确保这类系统的整体架构的安全，防止外部间谍活动和篡改。

第八章 拥护民主价值观：人工智能用于国家安全用途时的隐私、公民自由和公民权利

人工智能治理的民主模式

投资和使用人工智能工具来加强监督和审计

增加公共透明度

开发和测试用于保护隐私和公平性的系统

加强监督机制

保护法律补救措施和正当程序

美国政府的基本任务是保护美国人民的安全和自由。一直以来，美国民众对于“当安全和自由发生冲突时，如何才能实现‘鱼’和‘熊掌’可以兼得”这一问题始终争论不休。9/11 恐怖袭击之后的 20 年内，政府做了大量努力，将阻止另一次恐怖袭击的权力与尊重个人权利和自由的义务进行了协调。人工智能的出现，将这场辩论带入了一个全新的时代，因为新技术使得政府机构可以通过更加强大的方法收集和处理信息，跟踪个人行为和运动轨迹，并且根据计算机生成的分析结果采取行动。

除了支持海外军事行动和情报活动以外，人工智能工具在更接近美国本土的国家安全领域有着光明的应用前景——研究外国情报，寻找对美国构成危险的迹象；在边境筛查潜在威胁，防止网络攻击和信息战，识别国内的恐怖主义阴谋。美国民众对于人工智能应用在国家安全和公共安全领域的用途——特别是那些涉及生物识别技术或分析汇总的个人数据的应用——已经产生了巨大关切，认为这些技术将会侵犯他们的隐私，限制他们言论和集会的自由，增加偏见和歧视。与此同时，如果人工智能应用得当的话，它能提高对隐私和公民自由的保护。机器分析结果将会更加准确，而人工智能系统有可能通过实时监控加强监督。

对于美国和其他民主国家而言，政府官员在使用人工智能时，必须符合“有限政府”和“个人自由”原则。在民主社会中，任何赋予国家权力的行为必须伴随着明智的限制，从而确保这些行为在公民眼中合法化。

正如本报告所认为的，新兴人工智能技术在加强国家安全方面有着真实而重要的用途。美国情报、国土安全和执法部门出于国家安全的目的、开发和使用人工智能技术的能力必须得到维护。然而，要做到这一点，政府必须确保对于这项技术的使用是有效的、合理的以及合法的。公众的信任将取决于对隐私、公民自由和公民权利的合理保证。

民主的人工智能治理和隐私、公民自由和公民权利面临的全新挑战

随着新的技术专制治理模式在国外日渐得到青睐，美国必须继续充当民主价值观的灯塔。民主模式必须证明，在面对可能对其发起挑战的新兴技术变革时，民主模式本身具有强大的复原能力。从根本上说，我们相信美国的民主制度以及维护它的规则、规范和习俗，能够满足在人工智能时代维护安全和自由的双重需求。

对于情报部门而言，美国制度的核心特色包括最大限度减少收集、留存和传播美国公民数据相关的法律、规则和程序，这项制度的核心特色还包括情报部门必须接受政府三大机构的监管[1]。同样，国土安全和执法部门也必须在指导边境保护和犯罪调查的政策、监管和司法审查的框架内开展工作。归根结底，所有联邦机构的行动都受到宪法的约束。

在这种背景下，现代人工智能的进步以及它给情报、国家安全和执法任务带来的全新能力，引发了一系列关于美国隐私、公民自由和公民权利的关切、难题和挑战。例如：

- 人工智能驱动的分析可以协助政府官员处理和解读大量信息。通过汇总，这些信息可以形成一个人的活动、行踪和行为模式的“马赛克”图[2]。对于分析师或调查人员而言，当他们将涉及地理位置、网络浏览、金融交易和其他数据源的不同数据流结合起来时，就很有可能获得新的洞见。因此，这项技术对识别威胁非常有用，但也引起了关于边境或执法搜查的适当范围和授权问题[3]。

- 事实上，私人企业持有大部分私人信息。不过，现代数字生活的这一现实已经引发了公众对两大宪法问题的思考：当个人向技术公司等第三方提供信息时，是否应当以及何时应当拥有信息“合理的隐私预期”；当情报、国土安全和执法部门出于合法的国家安全目的，在何种情况下它们可以访问并利用我们的个人信息[4]。

- 人工智能有助于实现数据收集和分析的自动化。这种方法可以增强分析师和调查人员筛选和分流海量信息的能力，有助于其建立模式或定位威胁。但是，这也带来了关于机器和人类分析在这些流程中的适当作用问题，其中包括将数据用于预测性分析的问题。在某种程度上，人工智能系统的功能是不透明的，我们可能很难跟踪和证明导致系统做出推荐的计算过程。确定何时以及如何依靠算法，与情报部门最小化、开展询问程序以及为执法行动立案极其相关[5]。

- 基于不断变化的数据以及与其他模型的互动，人工智能模型可以不断演化，导致出现不可预期的后果。因此，与前几代技术相比，人工智能系统需要持续测试和评估。

- 在机器学习流程的很多阶段，可能无意间引入偏见，进而导致美国社会出现不平等现象，这一问题在执法环境中已有记载[6]。

管理人工智能挑战的原则

针对特定条件下，可能允许的或明智的行为，委员会不对此划定所有的界限。然而，在不同的国家安全背景下，应遵循以下几项重要原则：

外国情报收集与分析：国家情报总监办公室在发布的《情报部门人工智能道德指南》中，重点强调以维护美国民众隐私和公民自由的方式，利用人工智能获取海外情报，因此它被普遍看作是一项鼓舞人心的重大举措[7]。随着这些指导原则的实施，必须密切关注数据最小化、数据留存和问询程序已经得到充分和严格的执行。

边境安全：人工智能监视和分析能力可以使政府部门和机构在边境和入境口岸的执法工作变得更加高效。但是，为了维持民众对于人工智能在边境安全领域广泛应用的支持，国土安全部必须谨慎行事，确保自动筛查流程只会引导特工获得必要的和被授权可以获得的信息，并且确保特工不会在未获授权和许可的情况下，根据种族和宗教等特征，对个人进行筛选。

国内安全和公共安全：用于执法目的的人工智能技术，包括面部识别等生物特征监视技术，已经取得了快速发展，从而可能超过了人工智能技术的适当使用原则。因此，政府必须特别谨慎地管理与宪法基本原则相关的风险，包括平等保护、正当程序、免于无理搜查和扣押、言论自由和集会自由等[8]。

在执行任务时，在不同的国家安全背景下，保持政府主管部门之间的明显区分十分重要。同样，通过增加透明度、提升人工智能技术的性能和可靠性、确保正当程序和加强监督等措施，增强公众的信心，也十分必要。在牢记上述原则的基础上，政府应该采取下列步骤：

“政府机构应该评估人工智能应用方面的短期机会和研究空白，从而有效应对在隐私和公民自由上面临的挑战……”

投资和使用人工智能工具来加强监督和审计，支持隐私和公民自由。 政府机构应该评估人工智能应用方面的短期机会和研究空白，从而有效应对在隐私和公民自由上面临的挑战，例如：用于分类、推荐、异常检测和其他用途的机器学习技术[9]。人工智能在改善审计方面取得进展的例子包括支持财务审计和模型风险管理的工具。政府机构应该审查当前实践或新兴实践的实用性[10]。

提高政府使用人工智能的透明度。政府机构的政策和程序以及可能影响公民自由的人工智能的准确性都缺乏透明度[11]。部分机器学习系统的“黑匣子”性质也增加了这种不透明性[12]。提高透明度有助于缓解公众焦虑。当然，在特定的操作环境下，特别是对情报和执法部门而言，保密是执行任务的大前提。但是，我们可以更加有效地利用目前的透明机制，并在某些情况下，对透明机制进行改革。此外，全新的政府机构报告要求也有助于提高透明度。

- 对于影响美国人民的人工智能系统，国会应该要求情报部门、国土安全部和联邦调查局分别提交《人工智能风险评估报告》和《人工智能影响评估报告》。这些报告应该评估每个合格的人工智能系统或重大系统更新对隐私、公民自由和公民权利造成的影响[13]。

- 国土安全部和联邦调查局还应该改进发布记录系统通知和隐私影响评估的做法，以便在人工智能系统部署前，公众能够对其作用有一个更加全面的认识。

建议

开发和测试系统，推进隐私保护和公平性。特别值得注意的是，机器学习系统要求对隐私和公平性保证持续进行评估，包括对公平性的具体定义进行假设。虽然机器学习系统可能在一个静态时间点上满足要求，但一旦系统进入运行状态，持续的合规性就不再是必然的结果。这在很大程度上是由于不断变化的数据，引入非故意的偏见，以及潜在的匿名数据的重新识别[14]。这是一个复杂的技术领域，需要在技术、法律和政策领域持续进行工作，以便在保护隐私、公民自由和公民权利的技术方法上找到更多共识[15]。与此同时，政府机构应该采取下列步骤：

- **评估和降低人工智能系统设计、开发和测试风险。**除了对美国人民的隐私、公民自由和公民权利面临的风险进行评估以外，情报部门、国土安全部和联邦调查局应该采取措施来降低上述风险，并对可接受的剩余风险进行记录。通过这种做法，三家机构应该采用委员会在《以负责任的态度开发和部署人工智能系统时的关键考量因素》中推荐的一些做法，包括使用强大的匿名化等隐私保护手段，以及在条件允许的情况下，使用隐私保护技术；采取措施降低开发和测试中的偏见；持续对模型性能进行评估[16]。

- **针对可能影响隐私、公民权利和公民自由的人工智能技术，在部署之前，应在每一家政府机构中，指定办公室、委员会或团队对其进行审查。**这应该包括部署前的审查以及在系统生命周期内对合格性的审查。情报部门、国土安全部和联邦调查局的每一个办公室应具备评估数据、模型和系统记录的能力，对与系统预期用途相关的测试结果进行评估。

- **针对与国家安全有关且有可能影响美国民众的人工智能系统，组建第三方测试中心。**此类独立的第三方检测机构可以是国家实验室、高校附属研究中心或联邦政府资助研发中心。对于具有高度利益攸关性的系统而言，这种测试应该是强制性的；除此之外，此类测试是自愿性的[17]。测试将为政府机构提供额外的专业知识来克服内部限制。

建议

*针对受与人工智能相关的政府行为影响的人士，加强其寻求法律补救措施和正当程序的能力。*人工智能会犯错误[18]。政府机构在部署人工智能系统时，必须接受可能出现的“误报率”和“错报率”。与此同时，确保受与人工智能相关的政府行为影响的人士，例如由于系统错误，导致福利申请被拒绝（例如申请签证）、限制行动（例如被列入禁非名单）或逮捕[19]，能够获得符合宪法规定的正当程序的补救机会。在人工智能协助刑事指控案件的情况下，公众也存在着正当程序方面的顾虑[20]。我们建议政府机构采取两个步骤来解决这一问题：

- **审查国土安全部和联邦调查局内部可能影响正当程序，以及受影响人士寻求法律补救措施能力的政策和做法。**国土安全部和联邦调查局应该审查内部的政策和做法，确保受与人工智能相关的政府行为（包括系统行动和滥用）影响的各方，能够寻求法律补救措施并且清楚了解具体的办理流程。这种审查应该包括是否向受影响的各方提供了关于在决策中使用人工智能的充分告知，以及人工智能系统的审核程度，以便在有争议的情况下，可以追踪系统当时做出建议行动的全部过程。

- **发布《司法部长关于人工智能和正当程序的指导意见》**。指导意见应该说明在使用人工智能可能导致剥夺生命或自由的情况下，相关机构应如何保障美国人民的正当程序权利。

加强监督机制，解决当前和不断变化的问题。人工智能的发展需要一种前瞻性的监督方法来预测新技术的持续发展和应用，并使政府能够更好地在未来负责任地管理使用问题。

“人工智能的发展需要一种前瞻性的监督方法来预测新技术的持续发展和应用，并使政府能够更好地在未来负责任地管理使用问题。”

政府应该：

- **建立评估人工智能和新兴技术对隐私、公民自由和公民权利造成影响的特别工作组。**工作组的任务主要是发现差距，提出建议，确保人工智能和相关数据在美国政府行动中的使用符合美国法律和价值观，研究支持这一目标的组织改革。具体而言，工作组应该评估适用于当前人工智能应用和新兴技术的现有政策和法律法规上存在的空白，并针对下列领域给出下列建议：

    - 对人工智能和新兴技术及相关数据的开发和使用进行立法和监管改革[21]；

- 针对人工智能和新兴技术对隐私、公民自由和公民权利造成的影响，进行持续评估和经常性指导。

- **加强隐私与公民自由监督委员会的能力，使其能够对用于国家安全用途的人工智能进行有意义的监管并给出建议。**自 2007 年根据 "9-11 "事件委员会的建议成立以来，委员会在监督美国的反恐任务并向美国政府提供相关建议方面发挥了特别重要的作用。近年来，它开始将工作重心转向新技术在外国情报搜集和分析过程中起到的作用[22]。在人工智能系统部署之前，委员会应该被授权对人工智能系统的可见性，包括在更加细化的层面对其进行审视。委员会还应获得相应的资源和人员配置，以便完成人工智能时代需要的更加复杂的技术任务[23]。

- **对国土安全部隐私、公民权利和公民自由办公室进行赋权。**办公室主任必须与隐私官员进行协调，在采购和使用人工智能系统过程中，包括在国土安全部的机器学习系统中，使用相关数据的法律和批准程序中发挥不可或缺的作用。

- **要求联邦监督和审计组织之间加强协调和配合**。政府机构对人工智能文件和测试要求的遵守应得到严格的、技术上知情的监督支持。为实现这一目标并且克服目前的审计和监督障碍，应对一家常设机构进行调整和协调，以便加强人工智能在隐私、公民自由和公民权利方面的监督和审计[24]。

人工智能治理的民主模式

当人工智能用于国家安全用途时，人工智能治理应当满足当前和不断发展的隐私、公民自由和公民权利的需求

提高公共透明度。

国会应该要求情报部门、国土安全部和联邦调查局分别提交《人工智能风险评估报告》和《人工智能影响评估报告》。这些报告应该评估每个合格的人工智能系统或重大系统更新对隐私、公民自由和公民权利造成的影响。国土安全部和联邦调查局还应该改进发布记录系统通知和隐私影响评估的做法，以便在人工智能系统部署前，公众能够对其作用有一个更加全面的认识。

投资和使用人工智能工具来加强监督和审计。政府机构应该评估人工智能应用方面的短期机会和研究空白，从而有效应对在隐私和公民自由上面临的挑战，并且通过研究人工智能工具的发展，提升审计工作。

开发和测试系统，推进隐私保护和公平性。

政府机构应该评估和降低人工智能系统设计、开发和测试的风险。针对可能影响隐私、公民权利和公民自由的人工智能技术，在部署之前，应在每一家政府机构中，指定办公室、委员会或团队对其进行审查。针对与国家安全有关且有可能影响美国民众的人工智能系统，国会应当组建第三方测试中心

加强监督机制

政府应当组建评估人工智能和新兴技术对隐私、公民自由和公民权利造成的影响的特别工作组；加强隐私与公民自由监督委员会的能力，使其能够对用于国家安全用途的人工智能进行监管并给出建议；对国土安全部隐私、公民权利和公民自由办公室进行赋权；要求联邦监督和审计组织之间加强协调和配合。

保护法律补救措施和正当程序

国土安全部和联邦调查局应该审查内部可能影响正当程序原则及受影响人士寻求法律补救措施能力的政策和做法。司法部长应当发布《关于人工智能和正当程序的指导意见》。

第八章 – 尾注

[1]有关情报机构组成部分的总检察长指南汇编，请参见 E.O.12333，ODNI（2016 年 7 月 14 日）下总检察长批准的美国人程序的状态，https://www.dni.gov/files/documents/ Table_of_ EO12333 _ AG_Guidelines%20for%20PCLOB_%20Updated%20 July_ 2016.pdf.美国情报机构监督系统的组成部分包括情报机构、司法部、隐私和公民自由监督委员会等独立机构、联邦法院，包括外国情报监督法院以及众议院和参议院情报委员会的顾问和隐私官员。

[2]关于马赛克概念，请参阅斯蒂夫 • M • 贝劳文（Steven M. Bellovin）等人，*如果足够：位置跟踪，马赛克理论和机器学习*，纽约大学法律与自由杂志，第 8 卷（2013 年 9 月 3 日），https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2320019。

[3]关于《第四修正案》中关于人工智能背景下政府搜查条例的辩论，请参阅詹姆斯 E 贝克（James E. Baker），《半人马的困境：为即将到来的人工智能革命制定国家安全法》，第 6 章（布鲁金斯学会，2020 年）。

[4]国会和司法部门将需要评估目前对联邦政府获取和使用第三方数据（包括从数据经纪人处获得的数据）的法律约束是否充分。通过不断发展的判例法或立法，各机构将从第四修正案对第三方数据的适用方面的明确性中受益。关于第三方原则，见理查德 • M • 汤普森二世（Richard M. Thompson II），《第四修正案第三方学说》，国会研究局（2014 年 6 月 5 日），https://fas.org/sgp/crs/misc/R43586.pdf。

[5]有关智能环境下人类和机器分析含义的讨论和不同观点，请参见 Robert Litt，《信息时代的第四修正案》，《耶鲁法律杂志》（2016 年 4 月 27 日），https://www.yalelawjournal.org/forum/fourth-amendment-information-age； Cindy Cohn，信息时代保护第四修正案：对罗伯特 • 利特（*Robert Litt*）的回应，耶鲁法律杂志（2016 年 7 月 27 日），https://www.yalelawjournal.org/forum/protecting-the-fourth-amendment-in-the- information-age。

[6]对于算法错误率以及不同年龄、肤色和性别的不同表现，面部识别尤其令人担忧。请参阅帕特里克 • 格罗瑟（Patrick Grother）等人*的面部识别供应商测试，第 3 部分*：人口效应，国家标准与技术研究院（2019 年 12 月），https://nvlpubs.nist.gov/ nistpubs/ir/2019/NIST.IR.8280.pdf。性别阴影项目发现，各种面部识别系统对白人男性来说非常准确，但对女性和有色人种来说，它们的准确性明显较低（对有色人种女性来说最差）。参见性别阴影（上次访问时间：2021 年 1 月 11 日），http://gendershades.org/。

[7]请参阅*《智能社区人工智能道德的原则》*，国家情报总监办公室（上次访问时间：2021 年 1 月 11 日），https://www.dni.gov/index.php/features/2763-principles-of-artificial-intelligence-ethics -for-the - intelligence - communit y。

[8]一些观察员发现，“寒蝉效应”影响了个人行使言论、结社和集会自由的程度。请参阅雷切尔 • 莱文森-沃尔德曼（Rachel Levinson-Waldman）等，隐藏在众目睽睽之下：《分析政府公共监控的第四修正案框架》，埃默里法律杂志第 66 卷（2017 年），https://scholarlycommons.law.emory.edu/elj/vol66/iss3/4/.

[9]唐旭宁（Xuning (Mike) Tang）和石艺华（Yihua Astle），*T 深度学习对异常检测的影响*，Law.com（2020 年 8 月 10 日），https://www.law.com/legaltechnews/2020/08/10/the-impact-of-deep-learning-on- anomaly-detection/.

[10]例如，见 Bernhard Babel 等人，《机器学习和人工智能的衍生》，麦肯锡公司（2019 年 2 月 19 日），https://www.mckinsey.com/business-functions/risk/our-insights/derisking- machine - learning - and - ar tificial- intelligence；Saqib Aziz 和 Michael Dowling，《风险管理的机器学习和人工智能》，第 33-50 页（2018 年 12 月 7 日），https://link.springer.com/ chapter/10.1007/978-3-030 -02330 -0_ 3.

[11]例如，美国海关和边境保护局（CBP）在机场使用面部识别时的信息披露并不一致，而且有人声称联邦调查局没有按照法律规定提供有关其下一代识别数据库和使用面部识别的信息。2020 年，美国政府问责办公局（GAO）发现，"CPB 的隐私声明--告知公众其对该技术的使用--并非总是最新，或者在使用该技术的机场或 CBP 的网站上提供。”*面部识别：*美国海关与边境保护局（CBP）和美国运输安全管理局（TSA）正在采取措施实施这些计划，

但美国海关与边境保护局应解决隐私和系统性能问题，美国政府问责局（GAO）（2020 年 9 月 2 日），https://www.gao.gov/products/GAO-20 -568；另请参见永久阵容：《美国无管制警察面部识别》，乔治敦隐私与技术法律中心（2016 年 10 月 18 日），https://www.perpetuallineup.org/.

[12]在 2018 年的一份报告中，美国政府问责局对建立算法的科技公司缺乏透明度以及“对系统的准确性测试有限”提出了担忧。*人工智能：*新兴机遇、挑战和影响，美国政府问责局（2018 年 3 月），https://www.gao.gov/assets/700/690910.pdf.

[13]委员会建议，本章所述的工作队以及所附的行动蓝图应就两个问题提供具有约束力的指导：第一，情报机构、国土安全部和联邦调查局应在何时编制和发布人工智能风险评估报告和人工智能影响声明；第二，“合格的人工智能系统或重大系统更新”应由什么构成。”

[14]例如，化名数据可以与其他数据链接，以发现手机所有者的身份。参见 Byron Tau 和 Michelle Hackman，《联邦机构在移民执法中使用手机定位数据》，《华尔街日报》（2020 年 2 月 7 日），https://www.wsj.com/articles/federal-agencies-use-cellphone-location-data-for-immigration-enforcement-11581078600?mod=breakingnews.

[15]见本报告的附录，其中载有人工智能国家安全委员会的《负责任地开发和应用人工智能的主要考虑因素》的节选版本。供进一步讨论以下方面的建议：(1)采用与保护隐私相一致的技术和操作政策，并减少不必要的偏见；(2)持续监测和评估人工智能系统的性能，参见《负责任的人工智能开发和应用的关键考虑因素》的“使系统和用途与美国的价值观和法治相一致”和“系统性能”两节：*扩展版本*，人工智能国家安全委员会（2021）（已提交委员会存档）。

[16] *《负责任的人工智能开发和应用的关键考虑因素》*，人工智能国家安全委员会（2020 年 7 月），https://www.nscai.gov/previous-reports/。

[17]为了向各机构提供关于何时应利用这种测试机制的指导，一个组织应建立关于要求各机构进行第三方测试的阈值的指导。这应该包括一个人工智能系统何时可能对隐私、公民自由和公民权利构成足够高的风险，从而引发第三方审计师的测试要求的标准。

[18]请参阅卡什米尔·希尔（Kashmir Hill）的*被算法冤枉的人*，《纽约时报》 (2020 年 6 月 24 日)，https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html。

[19]个人的正当程序权利，包括通知，是建立在《宪法》和阐述该权利的判例法之上的。我们建议的目的是帮助确保政府在不使用这些长期权利的情况下履行其职责。

[20]正当程序权利要求个人有能力对针对他们的决定提出有意义的质疑。在联邦刑事审判中，这包括政府解释不利的决定是如何达成的，因此可以提出异议。在人工智能辅助或人工智能使能的决定的情况下，某些人工智能技术将不太有利于正当程序。参见丹妮尔·济慈·西特伦（Danielle Keats Citron），*技术性适当程序*，《华盛顿大学法律评论》（2008），https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview; 瑞安·卡洛（Ryan Calo）和丹妮尔·济慈·西特伦（Danielle Keats Citron），自动化行政国家：*合法性危机*，埃默里法律杂志（Emory Law Journal）（2020 年 3 月 9 日），https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3553590.人工智能系统的预测、分类或建议在法庭上受到挑战的早期案例表明，被告在寻求行使其权利时遇到了巨大的障碍。请参阅*诉讼算法：对政府使用算法决策系统的质疑*，AI NOW 研究所（2018 年 9 月），https://ainowinstitute.org/litigatingalgorithms.pdf。还有一些开放性问题，包括与人工智能有关的联邦证据规则和刑事诉讼程序。例如，在法庭上接受人工智能证据的证据标准尚待制定。

[21]例如，生物识别技术的基准人工智能标准和政策指导，政府采购商业人工智能产品，以及联邦数据隐私标准。

[22]请参阅*项目*，隐私和公民自由监督委员会（上次访问时间：2021 年 1 月 9 日），https://www.pclob.gov/Projects。

[23]隐私和公民自由监督委员会（PCLOB）与多个监督组织合作进行监督。同样重要的是，隐私和公民自由监督委员会（PCLOB）和这些组织要更好地调整和协调，在隐私、公民自由和公民权利方面进行互补的人工智能监督和审计。

[24]有关障碍的示例，请参见有贺孝（Taka Ariga）和斯蒂芬 桑福德（Stephen Sanford），*A 代表问责制：《人工智能时代的监督问题》*，欧洲审计团期刊，88-91（2020 年 1 月），https://www.eca.europa.eu/Lists/ECADocuments/JOURNAL20_01/JOURNAL20_01.pdf；另见新闻稿，情报机构监察长办公室，*美国情报机构监察长发表人工智能声明*(2019 年 5 月 30 日)，https://www.dni.gov/files/ICIG/Documents/News/ICIG%20News/2019/May%2030%20 -%20AI/Press%20Release%20 -%20AI. PDF；迈克尔 • 阿特金森（Michael K. Atkinson），*半年报告：2018 年 10 月 – 2019 年 3 月*，情报局监察长办公室（2019 年），https://www.oversight.gov/sites/default/files/oig-sa- Reports/20190430_ICIG-SAR_Oct18-Mar19.pdf。

第二部分

第九章　竞争与合作战略

组织美国政府应对新兴技术的挑战

成立技术竞争委员会

构建国家技术战略

建立高层次美中科技对话

人工智能（AI）对世界的影响远超狭义的国家安全概念。人工智能技术的发展构成了战略竞争的新支柱，并提升了在现有支柱中竞争的高度。经济基础最为坚韧、生产力最强大的国家才能担当世界领袖的职责。经济基础越来越取决于创新经济的力度，而创新经济又取决于人工智能技术的进步。人工智能技术将推动关键的基础设施、商业、交通、健康、教育、金融市场、食品生产和环境可持续性的前进浪潮。

对于人工智能和相关技术的研究、开发和应用竞赛，战略竞争愈发激烈。美国政府必须对人工智能竞争保持开放态度，建立相关组织，确保在竞赛中获胜。对过去几十年里美国十分凑效的美式创新方式，如今也必须进行微调，需以人工智能和相关技术竞争为中心来应对新出现的美中对抗。要保持美国的创新领导者身份和在全球的地位，美国需要一个更强有力的、由政府引导的技术战略，综合各项促进和保护措施，并把人工智能领域的各项投资连成更大的相关新兴技术群[1]。

“对于人工智能和相关技术的研究、开发和应用竞赛，战略竞争愈发激烈。美国政府必须对人工智能竞争保持开放态度，建立相关组织，确保在竞赛中获胜。”

本章阐述了人工智能竞争的本质以及要在竞争中获胜的两个先决条件：在白宫领导下成立技术竞争组织，建立与竞争对手的持续合作原则。后续章节（第 10 至 16 章）列举了整体战略的核心要素，描述了要确保美国在人工智能竞争中获胜须采取的行动，为赢得更广泛的技术竞争而奠定基础。首要的明确议程就是培养和招聘人才，推动形成一个多元化的人工智能创新生态系统，投入研发以利用人工智能和相关技术建设一个更为健康、繁荣和安定的社会。要更加激发国内的人工智能竞争力，并与世界上志同道合者联盟，对研究、知识产权（IP）和投资进行保护就显得必不可少。

美国政府必须正确解读技术竞争，并作出相应规定，成立有关组织，制定好相关条款，以应对中国的挑战。

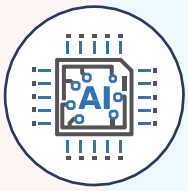
理解竞争

成立竞争组织

成立白宫国家技术竞争委员会：在白宫内单独授权一个部门，以设定战略方向，并监督和协调技术竞争。

中国制造 2025 和 2030AI 世界领导者：

中国已经构建了技术领导战略，选取了关键技术领域，在关键领域启动了高科技技术，并成立了跨政府机构的组织。



管理竞争

开启美中全面科技对话：建立与中国的高层对话，探讨面临的挑战，管理好新兴技术（例如 AI、量子技术、生物技术）有关的紧张形势。

赢得竞争

构建全国技术战略：技术竞争委员会应构建一个国家战略，以 AI 为出发点，引导新兴技术群的美国政策。

美中的人工智能竞争既重要又复杂

按照主要指标来评估美中的人工智能技术发展水平，通常是美国领先于中国[2]。然而，双方的差距在快速缩小。未来 10 年里，中国将顺理成章地超越美国，成为人工智能创新的中心[3]。近些年里，中国的科技公司已经在自然语言处理[4]、人脸识别技术[5]和其他人工智能技术应用领域中取得了突破性的进步。对于全球人工智能技术发展而言，中国企业、投资者、技术人员和学者都不可或缺。中国的社交媒体和电子商务企业在全球争夺用户。中国的电信企业在建设全球的 5G 基础设施。中国的风险投资家和大型科技公司大笔资金投入到初创公司中[6]。其先进的人工智能公司在美国[7]等地[8]拥有研究实验室。中国学者发表了一大批影响深远的的、推动着该领域进步的人工智能论文[9]。从国家安全角度而言，所有这些都不会构成美国的担忧，但中国的执政体制对美国利益构成了威胁。

“未来 10 年里，中国将顺理成章超越美国，成为人工智能创新的中心”

在一系列政府部门、大学和公司的人工智能规划导向下，中国的人工智能技术在飞速进步，其人工智能技术发展决心也更甚于美国[10]。一些战略性文件可反映出北京的观点，他们认为人工智能技术的进步在未来数 10 年中，将根本性重塑军事和经济的竞争态势[11]。最前沿的人工智能技术研究领域的科技公司和学术机构获得了中国政府的大额国家补贴支持，中国以此支持其战略规划发展[12]。中国利用西方完成的基础研究节省了资金，从而更聚焦在应用领域。中国在相关领域的研究和人才方面投入了大量资金[13]，并推动其“全国冠军级”企业去拓展海外市场[14]。通过其军民融合项目，中国寻求将商用和学术领域的人工智能技术进步融合到军事力量中[15]。利用技术转移项目和针对性投资，北京力求从美国和其他国家获取知识产权和数据[16]。

复杂的供应链网络、研究伙伴关系和企业关系将美中这两家世界上最大的人工智能领导者连接起来，也让美中竞争形势更加复杂化。采取极端措施来切断这些关联，对于美国而言代价高昂，并会造成全世界的强烈凡响。美中在学术、创新人才和市场之间的关系纠葛颇深，通常是互惠互利的关系，这也有助于推动人工智能技术领域的进步[17]。此外，利用官方外交展开人工智能技术和其它新兴技术上对话，探索在有益于人类的人工智能项目上进行合作的可能性，是符合美国国家利益的行动。

“美国可以开展和中国的竞争，而无须中断与中国的人工智能的合作研究，无须切断所有技术贸易。”

美国可以展开与中国的竞争，而无须中断与中国的人工智能合作研究，无须切断所有技术贸易。与中国的广泛技术脱钩可能会让美国大学和公司丧失稀缺的人工智能科学、技术、工程和数学（STEM）人才[18]，切断美国公司的有效供应链[19]，并切断市场及创新型公司的资金[20]。相反，针对性的脱钩只能作为美国整体方法的一个组成部分，如果能将之明智而谨慎地应用在关键核心领域，则可以帮助美国减少因非法技术转移引发的威胁，并保护核心的国家安全领域。

政策的挑战

中国的竞争方式不应该定义美国的创新方式，但中国确实展现了一种人工智能技术发展的的可替代模式，揭示了众多美国政府必须采取的公共政策选项以维护自身的优势。

美国需要重新检视自己的移民政策，以确保美国能打赢这场人工智能人才的竞争。美国需要以发展全球竞争力的视角审视各类人工智能和更广阔的 STEM 教育举措。出于计算和数据成本的考量，人工智能呈现了地域集中性和从大学转移到私营部门的趋势，因此美国必须考虑如何将人工智能研究进程多元化，并拓宽开展人工智能研究所需的数据和工具。在这个知识产权窃取随处发生的时代，美国必须考虑长久以来对于知识产权的处理方式是否恰当，而目前美国的知识产权体制在人工智能和其他新兴技术未能充分发挥作用。美国应保护其在微电子设计的领导地位，包括鼓励那些国家安

全所依赖的关键生产环节回流到国内。同时，美国必须确保充分利用防止违法技术转移的工具和政策，以应对军民两用新兴技术引发的国家安全挑战。

这些与人工智能有关的挑战逐渐暴露了更基本的问题，涵盖技术、经济和国家安全领域：

- 如何与对手竞争，而不牺牲美国价值观，包括自由市场原则、个人自由和有限监管。
- 如何确保国防和优先发展经济之间的适当平衡。
- 如何保持硬件优势，而不扼杀依赖外国竞争对手市场的国内设计人员和生产者。
- 出于国家安全目的考虑，如何利用和重塑私营部门的发展，而不扼杀私营部门导向的、自由市场的创新。
- 如何吸收最好的全球人才，而不让颠覆性技术和知识转移给竞争者。
- 如何培养一种开放的合作研究环境，同时填补被外国竞争对手利用的合法或非法漏洞。
- 如何维持长期的技术研发战略，应对快速变迁的地缘政治和技术发展。
- 如何保障投资/资本的自由流动，而避免战略竞争对手获得战略优势。
- 如何联合我方盟友和其他合作伙伴，减少其对中国数字技术的依赖性，建立更为坚韧的供应链，开发可以反映民主价值观的技术标准和规范。

在技术战略中需要一个更强有力的政府角色

委员会的目的不是要一个由国家掌控的经济，一个 5 年计划，或中国式的“军民融合”计划。而是要敦促由政府来主导的恢复政府、行业和学术界之间的均衡状态，以保障一个多元化的研究环境、有竞争力的经济，并符合国家需要的研究计划。当美国受到挑战时，通过调动行业和学术界并进行巨额投资，美国政府拥有丰富经验[21]。在面对一个像中国这样的竞争对手背景下，鉴于人工智能技术的变革潜力，美国将再次面对这一选择。

> “... 美国政府在发言和备忘录中表示：支持美国在人工智能技术中享有领导地位，但在实际商业投资和资助上投入有限...”

如今，美国政府在发言和备忘录中表示：支持美国在人工智能技术中享有领导地位，但在实际商业投资和资助上投入有限，并依赖于一个去中心化的治理结构去实现这点[22]。美国政府会谈论全球人才的竞争，但近年来，美国收紧了对高技术工作者的签证[23]。而美国从幼儿园到 12 年级的学生，在衡量 STEM 领域能力的考试中，已落后于东亚和欧洲的竞争对手[24]。技术领导者和政府认为“公私合作”很重要，但没有以具体的方式深入采取行动。美国专家们对人工智能技术霸权所带来的危险提出警告[25]，但华盛顿并没有构建新的联盟制定更民主的替代方案。当前的政策几乎就是对联邦政府中各种人工智能活动的汇编。然而，要赢得人工智能竞争胜利并保持美国的人工智能领导地位所需的战略却看不到。

政府必须精心筹划各类政策以推动创新，保护对国家安全至关重要的行业和部门，招募和培养人才，鼓励对国家安全和经济繁荣所必要的一系列技术进行研究，维持盟国和合作伙伴的民主规范。一些国家层面的战略需进行协调，并通过特定国家战略来支持人工智能研究、商业和教育。这需对促进和保护行为进行复杂的排序组合，以最大限度降低惩罚措施的成本和风险，确保基本研究和应用研究相互促进，协调与国际合作伙伴的合作方式，以及与立法机构核定执行优先级。这需要确认技术的趋势，评估美国和竞争对手的相对优点。总之，需要强有力和持续的白宫领导。

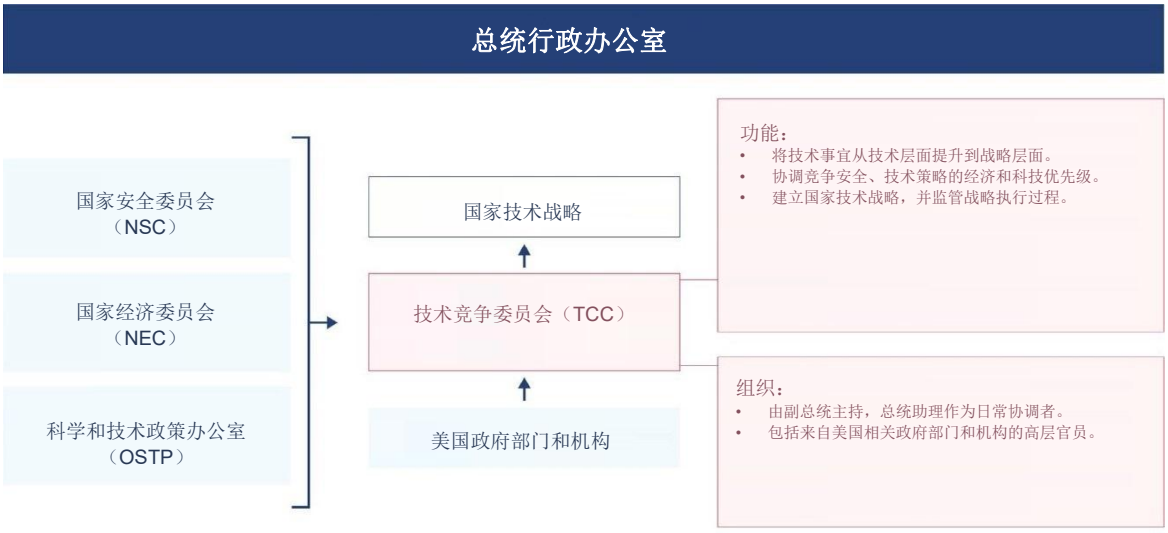
白宫领导的情况

政府需要一个权力中心，可以施展其领导力，推动国内经济、国家安全和科技政策，至今还没有这样的组织。总统行政办公室（EOP）有几个部门具有满足基本组织要求的部分职责和能力：国家安全委员会（NSC）[26]、科学和技术政策办公室（OSTP）[27]及其相关的国家科学和技术委员会（NSTC）[28]，以及国家经济委员会（NEC）[29]。国内政策委员会（DPC）起到关键的作用，在领导移民政策、教育政策和监管政策领

域具有职责[30]。还有一个部门——管理和预算办公室（OMB），负责审查相关预算和政府改革工作。

如果缺少这样重要的组织，就得由总统或副总统来确认、裁决和协调这些跨部门的事宜，此时会为别有用心者留下巨大的空间，从而影响总统决策。总统需要这样的组织，可以帮助其决策，通过现有的各委员会自上而下推动技术战略，并贯彻到所有政府部门。白宫应该：

技术竞争委员会.



“政府需要一个权力中心，可以施展其引导力，推动国内经济、国家安全和科技政策，这正是我们如今所缺少的。”

建议

*成立一个技术竞争委员会。*美国必须加强白宫对技术政策的行政领导力，授权单独的部门执行全面技术战略。委员会提议成立一个新的技术竞争委员会（TCC），包括总统行政办公室的领导人和内阁部长，可以召集机构召开白宫研讨会。TCC 由副总统主持，新任命一名负责技术竞争的总统助理作为日常领导人。技术竞争委员用于确保填补 NEC、OSTP 和 NSC 各方职责之间留下的空白，并和 OMB 相关联。该委员会不会取代 NSC、NEC 或 OSTP 领导的 NSTC 组织，而是提供一个研讨会，用于协调竞争安全、经济和科技优先级的问题，并将技术政策和技术问题从技术层面提升到战略层面。要协调委员会的工作，有必要任命一个新的负责人——总统在技术竞争力方面的助手，负责确保涉及新兴技术的政策会获得总统的足够关注。

建议

*建立国家技术战略。*基于以上的建议讨论，TCC 应构建国家技术战略，以人工智能为起点，在所有关键新兴技术方面引导美国政策。国家技术战略的目标，是保障美国整体在技术上处于长期领先地位，尤其是对于国家安全和竞争力至关重要的技术。战略应在竞争策略和优先级之间做权衡，确认在哪些关键技术上是竞争对手试图赶超美国领导地位的，并为新兴技术采取针对性对政策措施。首先，技术战略应基于以下支撑要素：1）赢得人工智能人才竞争；2）促进美国人工智能技术创新；3）保持美国的人工智能技术优势；4）引领有利的国际人工智能技术秩序。

建议

*建立高层次美中全面科技对话。*美国应和中国建立一个定期、高层次的外交对话，此对话应利于美国人民和我们的盟国，推动中国遵守国际规范。对话应聚焦新兴技术面临的挑战，包括人工智能技术、生物技术以及双方约定的其他技术。对话应具有两大重要目标：

- 确定用以解决全球性挑战的新兴技术合作领域，例如气候变化和自然灾害救助；
- 提供一个研讨会，用于探讨人们目前对于新兴技术特定用途的一些疑虑，并在两国之间建立关系并构建问题处理的通道。

第 9 章——尾注

[1]虽然美国政府已经签发了一批文件，强调 AI 技术研究和开发的重要性，例如：特朗普总统对于 AI 技术的行政命令。美国依然缺乏一种全面的政府整体计划，以引导政策制定者、研究学者和企业的方向，构建更安全的美国未来。《美国人民的人工智能》，白宫（最后访问时间：2021 年 1 月 28 日），https://trumpwhitehouse.archives.gov/ai/。

[2]例如: Alexandra Mousavizadeh 等人所著《全球 AI 指数》，Tortoise 传媒（2019 年 12 月 3 日） Media (Dec. 3, 2019), https://www.tortoisemedia.com/2019/12/03/global-ai-index/; 参见 Jean François Gagné等人所著《全球 AI 人才报告 2020》（最近登录时间：2020 年 12 月 29 日）, https://jfgagne.ai/global-ai-talent-report-2020/; 参见《全球 AI 人才追踪者》，MacroPolo（最近登录时间：2020 年 12 月 29 日） (last accessed Dec. 29, 2020), https://macropolo.org/ digital-projects/the-global-ai-talent-tracker/; 参见 Jeffrey Ding 等人所著《*MERICS 网络研讨会：中国是 AI 超级大国？量化中国与美国和欧洲的 AI 进程对比*》，MERICS（2020 年 7 月 1 日），https://merics.org/en/video/merics-web-seminar-china-ai-superpower-quantifying-chinas-ai-progress- against-us - and - europe.

[3]参见 Audrey Cher 所著《超级大国马拉松：美国可在科技上暂时领先中国，但北京有实力赶上》，CNBC（2020 年 5 月 17 日），https://www.cnbc.com/2020/05/18/us-china-tech-race- beijing- has - strength -to - catch- up - with - us - lead.html; 参见 Graham Allison 和 Eric Schmidt 所著《中国正在推翻美国的 AI 霸业？》，贝尔弗科学与国际事务研究中心（2020 年 8 月），https:// www.belfercenter.org/publication/china-beating-us-ai-supremacy; 参见 Will Knight 所著《仅仅两年内中国就可能以最好的 AI 研究超越美国》，麻省理工科技评论（2019 年 3 月 13 日）， https://www.technologyreview.com/2019/03/13/136642/china-may-overtake-the-us-with-the-best-ai-research-in-just-two-years/.

[4]参见 Karen Hao 所著《三张图显示中国 AI 产业如何被三家公司支撑起来》，麻省理工科技评论（2019 年 1 月 22 日），(Jan. 22, 2019), https://www.technologyreview.com/2019/01/22/137760/the-future- of-chinas-ai-industry-is-in-the-hands-of-just-three-companies/.

[5]参见 James Kynge 和 Nian Liu 所著《从 AI 到人脸识别：中国如何在新技术中设定规则》，金融时报（2020 年 10 月 7 日）https://www.ft.com/content/188d86df- 6e82-47eb-a134-2e1e45c777b6.

[6]参见 Yusho Chao 所著《中国风险投资人再次相中初创企业》，Nikkei（2020 年 9 月 13 日） https://asia.nikkei.com/Business/Finance/Chinese-venture-capitalists-take-a-shine-to-startups-again. 也可参见《设想中国科技巨头数十亿美元的收购案》， CB Insights（2020 年 5 月 28 日） (May 28, 2020), https://www.cbinsights.com/research/bat-billion-dollar-acquisitions-infographic/.

[7]参见例如《一家中国科技巨头正在亚马逊的主场设立一家 AI 科研实验室》，CNBC（2017 年 5 月），https://www.cnbc.com/2017/05/02/tencent-ai-research-lab-seattle.html.

[8]参见例如 Saheli Roy Choudhury 所著《阿里巴巴在中国以外成立聚焦在 AI 技术的联合 AI 科研中心》，CNBC（2018 年 2 月 28 日），https://www.cnbc.com/2018/02/28/alibaba-sets-up-joint-a-i-research- lab - in - singapore.html.

[9]人工智能促进协会（AAAI）是运营时间最长的 AI 协会之一，2019 年中国在该协会上提交和发表论文数量最多。参见人工智能指标：2019 年年报，斯坦福以人为本人工智能研究院，41（2019），https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf.艾伦人工智能研究所也预测，中国有信心到 2025 年在人工智能的期刊论文最佳引文、突破性论文份额上超过美国。参见 Field Cady 和 Oren Etzioni 所著《中国可能在 AI 技术研究上超过美国》艾伦人工智能研究所（2019 年 3 月 13 日），https://medium.com/ai2-blog/china-to-overtake-us-in-ai-research-8b6b1fe30595.

[10]该战略文件的汇编，参见《AI 政策：中国》，未来生命研究所（最近登录时间：2020 年 12 月 30 日） https://futureoflife.org/ai-policy-china/; 参见 Graham Webster 等人所著《全文翻译：中国的“新一代人工智能发展规划”》，新美国（2017 年 8 月 1 日），https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new- generation-artificial-intelligence-development-plan-2017/

（翻译中国国务院 2018 年 7 月 20 日印发的新一代人工智能发展规划通知）。

[11]参见 Gregory C. Allen 所著《理解中国的 AI 战略》，新美国安全中心（2019 年 2 月 6 日），https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

[12]参见 Ashwin Acharya 和 Zachary Arnold 所著《中国公开 AI 研发开支：暂时发现》，安全和新兴技术中心（2019 年 12 月）,https://cset.georgetown.edu/wp-content/uploads/ Chinese-Public-AI-RD-Spending-Provisional-Findings-1.pdf; 也可参见 Emily Weinstein 所著《描绘中国招聘科学家的庞大工作》, 防务一号（2020 年 11 月 30 日），https://www.defenseone. com/ideas/2020/11/mapping-chinas-sprawling-efforts-recruit-scientists/170373/; 参见 David Cyranoski 所著《中国加入 AI 人才战》，自然杂志（2018 年 1 月 17 日），https://www.nature.com/articles/d41586 - 018- 00604-6.

[13]*同上。*

[14]美中经济与安全审查委员会，《关于技术、贸易和军民融合的听证会：中国寻求 AI 技术、新材料和新能源》，46、115-116（2019 年 6 月 7 日），https://www.uscc.gov/sites/default/files/2019 -10/June%207,%202019%20Hearing%20Transcript.pdf.

[15]美中经济与安全审查委员会，《关于技术、贸易和军民融合的听证会：中国寻求 AI 技术、新材料和新能源》（2019 年 6 月 7 日）https://www.uscc.gov/sites/default/files/2019 -10/June%207,%202019%20Hearing%20Transcript.pdf.

[16]*同上。*

[17]这是 Eric Schmidt 在新技术关系和竞争关系建立过程中所注意到的。参见 Hal Brands 和 Francis J. Gavin 所著《新冠病毒和全球秩序：冲突、竞争与合作的未来》，约翰霍布金斯大学出版社， 406-418（2020 年 8 月 31 日），https://muse.jhu.edu/chapter/2696578.

[18]参见 Ishan Banerjee 和 Matt Sheehan 所著《美国的 AI 人才获取：美国在 AI 研究的巨大领先优势是依靠输入研究人员》，MacroPolo（2020 年 6 月 9 日）， https://macropolo.org/americas-got-ai-talent-us- big-lead-in-ai-research-is-built-on-importing-researchers/?rp=m.

[19] *《美国对中国稀土的依赖：贸易战弱点》*， Reuters（2019 年 6 月 27 日），https://www.reuters.com/article/us-usa-trade-china-rareearth-explainer/u-s-dependence-on-chinas-rare- earth-trade-war-vulnerability-idUSKCN1TS3AQ.

[20]参见 *Member Survey* 所著《美中商业委员会》（2019 年 8 月），https://www.uschina.org/sites/default/files/member_survey_2019_-_en_0.pdf.

[21]例如，经通胀调整后，曼哈顿计划估计耗费 270 亿美元，阿波罗计划总共耗费约 1210 亿美元。 美国国会研究服务部，Deborah Stine 所著《曼哈顿计划、阿波罗计划和联邦能源技术研发项目的比较分析》（2009 年 6 月 30 日）https://fas.org/sgp/crs/misc/RL34645.pdf（使用美国劳工统计局的 CPI 通胀计算器转化为 2020 年美元金额，见 https://www.bls.gov/data/inflation_calculator.htm).

[22] 2018 年美国联邦研发投入占 GDP 的比例，从 1970 年代峰值时约 2%回落到 0.7%。参见对外关系委员会 James Manyika 和 William H. McRaven 所著《创新和国家安全：保持我们的优势》（2019 年 9 月），https://www.cfr.org/report/keeping-our-edge/ recommendations/.

[23]参见 Zolan Kanno-Youngs 和 Miriam Jordan 所著《特朗普收紧外国高新技术工作者签证》，纽约时报（2020 年 10 月 6 日），https://www.nytimes.com/2020/10/06/us/politics/h1b-visas- foreign-workers-trump.html.

[24]参见 Moriah Balingit 和 Andrew Van Dam 所著《美国学生在阅读、数学和科学考试中继续后于东亚和欧洲同龄人》，华盛顿邮报（2019 年 12 月 3 日），https:// www.washingtonpost.com/local/education/us-students-continue-to-lag-behind-peers-in-east-asia- and-europe-in-reading-math-and-science-exams-show/2019/12/02/e9e3b37c-153d-11ea-9110 - 3b34ce1d92b1_story.html.

[25]参见 Alina Polyakova 和 Chris Meserole 所著《输出数字权威主义》，Brookings（2019 年 8 月），https://www.brookings.edu/research/exporting-digital-authoritarianism/.

[26]国家安全委员会有一个法定职责是“就国家安全有关的国内、国外和军事综合政策为总统提供建议，确保美国军方和其他美国政府部门和机构在涉及国家安全的事件中合作更为有效”， 50 U.S.C. § 3021(b)(1) 。

[27]参见 Pub. L. 94-282，《国家科学和技术政策、组织和重点法案》， 1976、90 Stat. 459（1976 年），https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ ostp_organic_statute.pdf.

[28]在 OSTP 主任监督下的 NSTC 的职责是：“（1）协调科学和技术政策制定过程；（2）确保科学和技术政策决策和项目与总统的国家目标一致；（3）帮助整合总统的科学和技术在联邦政府的政策议程；（4）确保在联邦政策和项目的制定和执行中纳入科学和技术内容；（5）推动科技和技术的国际合作。助理可采取此类行动，包括为执行此行动所需或所适宜而起草宪章”。威廉·克林顿签发的 12881 号行政命令：成立国家科学和技术委员会（1993 年 11 月 23 日），https://www.govinfo.gov/content/pkg/ WCPD-1993-11-29/pdf/ WCPD-1993-11-29 - Pg2450.pdf.

[29]参见 William J. Clinton《行政命令 12835 号：成立国家经济委员会》（1993 年 1 月 5 日），https://www.govinfo.gov/content/pkg/ WCPD-1993- 02- 01/pdf/ WCPD-1993- 02- 01-Pg95.pdf.

[30]参见 William J. Clinton，《行政命令 12859 号：成立国内政策委员会》（1993 年 8 月 16 日）https://www.archives.gov/files/federal-register/executive-orders/pdf/12859.pdf.

第十章 人才竞争

赢得 AI 人才竞争

聚焦数字技能的第 2 版 NDEA

吸引并留住世界的聪明才俊

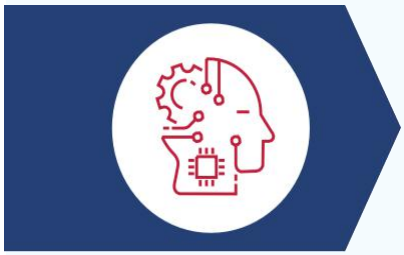
全球都在争夺稀缺的人工智能人才，美国也深陷此竞争之中[1]。委员会对当前的人才趋势忧心忡忡。加入人工智能博士项目的美国本土学生人数自从 1990 年后就未曾增长，而争夺国际学生的竞争却愈演愈烈，这危及到美国留住国际学生的能力[2]。美国自成立以来首次面临输掉科学前沿人才竞争的风险。要保持美国的领先地位，有两个选择：培养更多本土有潜力的人才，招募和留住更多现有的外国人才。

“美国自成立以来首次面临输掉科学前沿人才竞争的风险。”

无论是竞争者，还是我们的盟国，都认识到实施人工智能人才战略的重要性。从 2000 年到 2014 年，中国大学在科学、技术、工程和数学（STEM）领域的研究生数量增加了 360%，仅 2014 年就有 170 万人[3]。而同时期内，美国大学在 STEM 相关领域的研究生人数只增加了 54%，其中许多是国际学生[4]。如今中国研究人员数量约占全球顶尖的深度学习人才数量的 29%[5]。中国和其他国家也采取了吸引国际人才的措施，包括灵活的移民政策和为科技人才提供各项激励措施[6]。

美国想要在现在和未来留在人工智能技术队伍中的头部人才，就应在人工智能人才渠道上大举投入。否则，面对这种人工智能人才竞争态势，必定难有作为。

要实现在 AI 技术上的优势地位，美国需要培养以下四种典型人才，以推动美国的 AI 技术发展：研究人员、开发者、终端用户和知情消费者。



研究人员

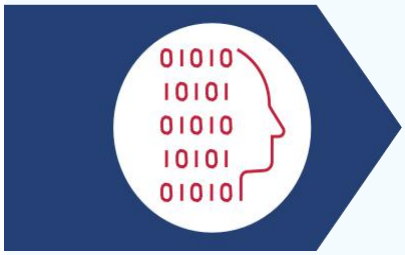
AI 研究人员将致力于半自动和全自动系统的实现和研究开发；作为算法专家，熟悉现代 AI 领域最前沿的研究成果，帮助形成 AI 领域新的理念；在一个重大项目中负责从基础研究到原型测试的开发。

开发者

负责数据清洗、特征抽取、选择及分析；模型训练和模型校正；与领域专家及终端用户合作；以及发现新的应用场景；开发者的培训和教育程度无须达到 AI 专家的水平，大致等同于专科或本科学位的培训、教育或经验即可，培训和教育也包括在数据处理和模型训练中如何处理相关道德问题及偏见问题。

终端用户

他们将 AI 技术用于其日常业务中。AI 的应用与使用现有软件非常类似，都需要一些系统特定的培训，仅有某些管理数据的职位无须具备 AI 特定的专门知识。



知情消费者

这个人群在采购技术时，有能力做出更好的消费者选择，并理解其市场行为的重要性。

拓展 STEM 的承诺和局限。

在 STEM 教育方面进行投资，是提高美国国家实力，改善国家安全的必然之举。虽然美国从幼儿园到 12 年级的教育体系质量表现参差不齐，但因为具有吸引国际人才的强大能力，美国在国际人才举措上整体依然排名居前[7]。美国在 STEM 教育上的大量投入至关重要，这是推动美国人工智能人才增长的引擎。然而，仅在 STEM 教育上的投入还不足以让美国赢得人工智能和 STEM 人才的国际竞争。中国正在产出大量的计算机科学家、工程师和其他 STEM 的研究生[8]。在可预见的未来，美国的 STEM 教育体系无论是在产量上还是质量上，都无法为美国市场或国家安全企业提供充足的 STEM 或人工智能人才[9]。要想竞争，美国必须改革其教育体系，更高质量、更大数量地产出研究生。

建议

*通过第 2 版国防教育法案。*1957 年，在苏联发射了“斯普特尼克”人造卫星后，出于在教育和创新方面落后于人的担心，美国国会在次年通过了国防教育法案（NDEA）。NDEA 强调了学生学习科学、数学和外语的重要性，批准超过 10 亿美元用于减少学生贷款、各级教育资金和研究生奖学金。许多学生因为该法案才得以读大学。1960 年，有 360 万学生上了大学，到 1970 年这个数字变为 750 万人[10]。这个法案帮助美国赢得了太空竞赛，推动了微电子行业，加速了美国的创新能力提升，并最终为美国赢得冷战胜利起到重要作用。

委员会相信当下是第 2 版 NDEA 出台的好时机，其参照了第 1 版的条款内容，但也具有重要差异。第 2 版 NDEA 将聚焦在资助学生获得数字技能，例如数学、计算机科学、信息科学、数据科学和统计学。第 2 版 NDEA 应包括从幼儿园到 12 年级的教育和再培训项目，有目的地针对资源不足的学校地区，解决美国教育体系全方位的不足。委员会也建议大力投入大学水平的 STEM 项目，为 25000 名大学本科生、5000 名研究生和 500 名博士提供奖学金。本科奖学金应包括社区大学的学生，以确保更多美国人可以获得负担得起的 STEM 教育。第 2 版 NDEA 的最终目的是通过激励未被充分代表的美国人民的项目，扩展数字人才库。

“委员会相信当下是第 2 版 NDEA 出台的好时机...”

建议

*通过移民壮大人工智能人才。*移民改革对于国家安全而言势在必行，能成功吸引并留住高技术人才的国家，将获得比竞争对手更大的战略和经济优势。人力资本优势在人工智能领域尤其突出，人工智能人才供不应求[11]。高技术移民可加速美国创新，推动创业并制造就业机会[12]。美国从高技术外国移民获得的好处远大于其他国家。2013 年，移民到美国的发明家数量为旅居海外的美国发明家数量的 15 倍[13]。相比之下，加拿大、德国和英国的发明家移民比例都是净负值[14]。与美国在人工智能领域的其他竞争优势相比，例如财力或硬件实力，这种移民优势是其他国家更加难以复制的。

“能成功吸引并留住高技术人才的国家，对竞争对手拥有战略上和经济上的优势。”

不幸的是，美国的国际学生越来越多地选择在其他国家学习或毕业后回国[15]。其中一个原因是积压的绿卡申请越来越多[16]。印度移民面临着超长等待期。许多人用几十年的工作签证来等待批准绿卡，既阻碍了技术部门招聘人才的能力，又影响了印度移民的生活质量。同时，其他国家不断在努力吸引和留住人工智能人才，包括那些在硅谷中的移民[17]。

“这些移民的限制，损害了美国的创新和经济增长，只会帮助我们的竞争对手实现人力资本的增长。这些原因也促使美国的技术公司搬迁到人才聚集之处，或是海外。”

虽然移民有益于美国，但政策制定者也必须牢记不利的技术转移所带来的威胁。然而，限制移民显然不是解决这个问题的方法[18]。这些对移民的限制，损害了美国的创新和经济增长，只会帮助我们的竞争对手实现人力资本的增长。这些原因也促使美国的技术公司搬迁到人才聚集之处，跨境或是海外[19]。如果美国技术领域主要的研究和开发转移到中国或其他比美国更易受技术转移影响的国家，那么技术转移问题只会更加恶化[20]。更有效的方式是在采取行动提升美国吸引全球顶级人才能力的同时，采取有针对性措施打击技术转移载体。NSCAI 在本报告第 14 章中详细阐述了技术转移的内容。其建议应配合移民政策改革一同进行。

中国政府认识到美国吸引和留住中国人才的能力，非常重视人才外流的威胁。移民政策同样减缓了中国人工智能技术的发展进程，构成了中国崛起的一道障碍[21]。中国人才日益外流，使得中国政府陷入了两难境地：一方面面临流失更多人力资本，进而降低经济的增长，危及人工智能领域进步，反过来，则是抹杀中国民众在美国学习和工作的机会。与此同时，美国也应谨慎本国人才流入合作伙伴国。

“增加人才外流将导致中国陷入两难境地…”

建议

*扩大“杰出”人才范围，让 O-1 签证更容易获取，并着重于人工智能人才。*O-1 临时工作签证是用于具有杰出能力或杰出贡献的人才。目前，签证官是通过主观评估来决定申请人的资格。对于科学和技术，这个评估方法和学术标准相当一致，例如在主要刊物上的论文发表数量，但却不适用于在实业中表现优秀的人。

建议

*执行并宣传国际创业者条例。*国际创业者条例（IER）允许美国公民及移民局（USCIS）批准国际创业者在美国合法逗留，只要他们可以证明“他们在美国逗留，能通过创业创造显著的公众利益”[22]。行政行动可公开宣称其管理目的是通过 IER 来推动移民创业者为美国创造就业以及推动经济增长。USCIS 也可以公开宣称，其将为诸如人工智能一类的高优先级 STEM 领域的创业者，或是人工智能在其他行业（如农业）应用领域的创业者，提供优先移民考虑。创业者吸引投资的能力也应作为筛选创业者的一个标准。

建议

*提升高技术工作者的工作可更换性，并做相应澄清。*持 H-1B 和 O-1 和其他临时工作签证的工作者，要获得一年可续签的开放式工作许可的标准限制太多，内容也不够清晰，应做出变更，澄清何时高技术、非移民工作者被允许更换工作或雇主，以便当雇主撤销申请或歇业时可以提高工作的灵活性，提高了 H-1B 签证工作者寻求其他 H-1B 工作的灵活性。

建议

*恢复发放因为官僚错误而流失的绿卡。*联邦政府机构签发绿卡的数量通常低于他们允许发放的数量。自 2009 年起，联邦政府因为累积的官僚错误而流失的绿卡数量超过 32.6 万张[23]。国务院和国土安全部（DHS）应该公布最新的因为官僚错误而流失的绿卡数量报告。两个部门都可以通过相关机构给排队等候的申请者发放流失的绿卡。国会应通过相关立法，支持国务院和国土安全部恢复发放流失的绿卡[24]。

*将绿卡授予官方认可的美国大学 STEM 博士毕业的学生。*国会应该修订移民和国籍法[25]，对于所有被审查过的（无国家安全风险）、在 STEM 相关领域以全日制或半脱产和远程授课混合方式，毕业于官方认可的美国高等院校获得博士学位，且在科学、技术、工程或数学有关领域找到工作的人员授予合法永久居留身份。他们不应计入永久居留身份的数量上限中。

建议

*职业移民绿卡的数量翻倍。*在现有体系下，职业移民绿卡数量过于稀少：每年 14 万张，其中不到一半提供给重要岗位[26]。这让许多高技术工作者无法获得永久居留权，并无法像本土工作者一样有效换工作或与雇主谈判，从而导致美国对高技术工作者的吸引力出现下降。要减少这种高技术工作者申请绿卡的积压情况，美国应该将职业移民绿卡的数量翻倍，并突出强调为 STEM 和人工智能相关领域高技术工作者授予永久居留身份。

建议

*提供创业者签证。*国际博士生比本土学生更愿意去创立一家公司，或去初创公司工作，但他们难以选择这种方式留在美国。原因之一是 H-1B 签证系统的限制[27]。类似，那些没有充足资金通过 EB-5 签证获得永久居留权的移民创业者，不得已只能使用其他针对在现有公司而非初创公司里工作的学者和工作者的签证来获取永久居留权[28]。所有这些问题都让美国对于国际人才的吸引力出现下降，也许更重要的是，这样降低了初创公司和其他小公司（他们是美国新就业机会的主要来源）雇佣高技术移民的能力，这些移民已经被证明了其提高公司成功几率的能力。国会应该提供一种新的创业者签证，用于那些可在有限试用期内发展壮大其公司，为美国提供“显著的公众利益”的人们[29]。这个签证是作为职业签证、投资签证或学生签证的替代方案，主要用于潜在的创业人群。

建议

*提供一种新兴和颠覆性技术签证。*国家科学基金会（NSF）应每三年确认一次关键的新兴技术清单。DHS 则允许在该适用领域中的学生、研究人员、创业者和技术人员申请新兴和颠覆性技术签证。这可以提供急需的研发人才，增强经济实力[30]。

第 10 章——尾注[31]

[1]对于填补 AI 人才缺口所需数量，各种评估差别很大，但各方的共识是，当国家为稀缺资源而竞争时，现在的人才缺口很明显，将来也会非常显著。参见 Remco Zwetsloot 等人所著《加强美国劳动力：政策和研究议程》，安全和新兴技术中心，（2019 年 9 月 2 日），https://cset.georgetown.edu/wp- content/uploads/CSET-Strengthening-the-U.S.-AI-Workforce.pdf （“腾讯研究院，一家主要的中国技术公司，声称全球有大约 30 万名 AI 研究人员和从业者，而市场需求达到数百万之多。Element AI，一家加拿大 AI 公司，在 2018 年估计全球有大约 22,000 名从事 AI 研究的博士级研究人员，其中仅有 25%的人精通技术，可以带领团队从研究走向应用。AI 公司 Diffbot 估计全球有超过 70 万名擅长机器学习的人员。”）

[2]参见 Remco Zwetsloot 等人所著《将顶尖 AI 人才留在美国》，安全和新兴技术中心，iii-vi（2019 年 12 月）。https://cset.georgetown.edu/wp-content/uploads/Keeping- Top-AI-Talent-in-the-United-States.pdf.

[3] 《中国在科学和工程上的崛起》，NSF 国家科学委员会（2018 年），https://www.nsf. gov/nsb/sei/one-pagers/China-2018.pdf （中国的同行评议的科学和工程论文占全球的份额也超过美国）。

[4] 《科学和工程指标 2018》，NSF 国家科学委员会（2018 年），https://www.nsf.gov/statistics/2018/nsb20181/assets/561/higher-education-in-science-and-engineering.pdf.

[5]为此，2019 年在久负盛名的 AI 深度学习会议神经信息处理系统会议上，会议录用的论文对“顶尖”人才做出定义。“顶尖”人才代表了该领域中顶尖的 20%研究人员。见 MacroPolo 的《全球 AI 人才追踪者》（最后登录时间：2020 年 12 月 28 日）https://macropolo.org/digital-projects/the-global-ai-talent-tracker/。中国已在重点关注深度学习，这只是 AI 的许多重要方面之一。

[6]例如，中国的千名人才计划是国家组织的，旨在 2050 年成为全球科技领导者蓝图的一部分。工作人员报告，《对美国研究企业的威胁：中国人才招募计划》，美国参议院常设调查委员会（2019 年 11 月 14 日）https://www.hsgac.senate.gov/imo/media/doc/2019 -11-18%20PSI%20Staff%20Report%20 -%20China’s%20Talent%20Recruitment%20Plans.pdf.

[7] 《全球 AI 人才追踪者》，MacroPolo（最后登录时间：2020 年 12 月 28 日），https://macropolo.org/ digital-projects/the-global-ai-talent-tracker/. See also, *High- Skilled Immigration and the Rise of STEM Occupations in U.S. Employment*, National Bureau of Economic Research at 1 (Sept. 2016), 也可参见 Gordon Hanson 和 Matthew Slaughter 所著《高新技术移民和 STEM 职业在美国就业中的增长》，国家经济研究局，1（2016 年 9 月） https://www.nber.org/system/files/working_papers/w22623/ w22623.pdf.

[8] 《中国在科学和工程上的崛起》，NSF 国家科学委员会（2018 年），https://www.nsf. gov/nsb/sei/one-pagers/China-2018.pdf.

[9]正如本报告第 6 章所指出，2019 年美国有 433,116 个计算机科学职位，而 2019 年美国大学新毕业的计算机科学家只有 71,266 名。Code.org（最后登录时间：2021 年 1 月 11 日）https://code.org/promote。也参见 Oren Etzioni 所著《特朗普行政命令在 AI 技术上的缺失：美国需要特别签证项目，旨在吸引更多 AI 专家》，《连线杂志》（2019 年 2 月 13 日），https://www.wired.com/story/what-trumps- executive-order-on-ai-is-missing/。

[10] 《斯普特尼克卫星刺激美国通过国防教育法案》，美国参议院（最后登录时间：2020 年 12 月 28 日）https://www.senate.gov/artandhistory/history/minute/Sputnik_ Spurs_ Passage_of_ National_ Defense_ Education_ Act.htm#:~:text=The%20National%20Defense%20Education%20Act%20of%201958%20became%20one%20of,and%20private%20colleges%20and%20universities.

[11]根据一份报告，在一个流行的求职网站上，AI 的职位列表数量“从 2015 年到 2017 年增长超过 5 倍，深度学习技术需求增长超过 30 倍”，而在这个领域求职的人数则增长得慢多了。这种不匹配正在拖累 AI 技术的采用。大部分公司报告认为技能的缺口是阻止他们采用 AI 技术的主要障碍之一。参见 Remco Zwetsloot 等人所著《加强美国劳动力：政策和研究议程》，安全和新兴技术中心，（2019 年 9 月 1 日），https://cset.georgetown.edu/wp-content/uploads/CSET-Strengthening-the-U.S.-AI-Workforce.pdf.

[12]参见 William S. Kerr 所著《高新技术移民、创新和创业：实证方法和证据》，国家经济研究局，7-8（2013 年 8 月），https://www.nber.org/ papers/w19377; 参见 Gordon Hanson 和 Matthew Slaughter 所著《加强美国 AI 劳动力、高新技术移民和 STEM 职业在美国就业中的增长》, 国家经济研究局，23（2016 年 9 月），https://www.nber.org/system/files/working_papers/w22623/ w22623.pdf; 参见 Remco Zwetsloot 等人所著《加强美国 AI 劳动力：政策和研究议程》，安全和新兴技术中心，5（2019 年 9 月）， https://cset.georgetown.edu/

wp-content/uploads/CSET-Strengthening-the-U.S.-AI-Workforce.pdf.

[13]参见 Carsten Fink 所著《是什么让发明人迁移？》，世界经济论坛（2013 年 7 月 17 日），https://www.weforum.org/agenda/2013/07/what-leads-inventors-to-migrate/.

[14] 参见 Ernest Miguelez 和 Carsten Fink 所著《测量发明人的国际流动性：一个新数据库》，世界知识产权组织，16（2013 年 5 月）， https://www.wipo.int/edocs/pubdocs/en/ wipo_pub_econstat_wp_ 8.pdf.

[15]*根据安全和新兴技术中心的统计，2016 年国际学生中有 14%拒绝了在美国大学学习的录取通知书，改为在本国学习，19%的人决定在另一个国家学习。2018 年，这些数字上升为 39%在本国学习，59%去另一个国家学习。参见 Remco Zwetsloot 等人所著《将顶尖 AI 人才留在美国：留住国际研究生的调查结果和政策选项》，安全和新兴技术中心（2019 年 12 月 26 日）https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the- United-States.pdf.*

[16]参见 Shulamit Kahn 和 Megan MacGarvie 所著《STEM 博士永久居留申请延误的影响：谁离开，为什么》，研究策略（2020 年 11 月），https://www.sciencedirect.com/science/article/ abs/pii/S0048733319301982.

[17]参见 Tina Huang 和 Zachary Arnold 所著《移民政策和全球 AI 人才竞争》，安全和新兴技术中心，8（2020 年 6 月），https://cset.georgetown.edu/research/i mmigration-policy-and-the-global-competition-for-ai-talent/.

[18]参见 Zachary Arnold 等人所著《移民政策和美国 AI 领域：初步评估》，安全和新兴技术中心，22（2019 年 9 月）， https://cset.georgetown.edu/research/ immigration-policy-and-the-u-s-ai-sector/.

[19]参见 Remco Zwetsloot 等人所著《加强美国 AI 劳动力：政策和研究议程》，安全和新兴技术中心，5（2019 年 9 月）， https://cset.georgetown.edu/wp- content/uploads/CSET-Strengthening-the-U.S.-AI-Workforce.pdf.

[20]中国是世界上最大的 AI 人才单一来源，领先的美国技术公司，例如谷歌和微软都已在中国建立了前沿科技研究中心，部分原因是为了获取人才。这增加了中国的 AI 研发实力和技术转移的可能，并且如果公司留在美国，就减轻了美国情报体系（IC）对收集中国技术发展信息的合法授权。参见 MacroPolo 所著《全球 AI 人才追踪者》（最后登录时间：2020 年 1 月 17 日）https://macropolo.org/digital-projects/the-global-ai-talent-tracker/; Roxanne Heston & Remco Zwetsloot, *Mapping U.S. Multinationals' Global AI R&D Activity*, Center for Security and Emerging Technology at 20 (Dec. 2020), https://cset.georgetown. edu/wp-content/uploads/CSET-Mapping-U.S.-Multinationals-Global-AI-RD-Activity-1.pdf.

[21]参见 Remco Zwetsloot 所著《美中 STEM 人才"脱钩"：背景、政策和影响》，约翰霍普金斯应用物理图书馆，19（2020 年），https://www.jhuapl.edu/assessing-us-china- technology-connections/dist/407b0211ec4929960855132604148 8d4.pdf（中国共产党中央人才工作协调小组组长……抱怨"中国流失的顶级人才数量世界排名第一。"）也参见 Joy Dantong Ma 所著《中国 AI 人才基础正在增长并离开》，MacroPolo（2019 年 7 月 30 日） https://macropolo.org/ chinas-ai-talent-base-is-growing-and-then-leaving/?rp=m（注意：2009 年至 2018 年，参加过 NeurIPS 会议的 2800 名中国与会人员中，有四分之三在中国以外工作）。

[22]《外籍创业者特殊工作许可项目》，USCIS（2018 年 5 月 25 日），https://www.uscis.gov/humanitarian/humanitarian-parole/international-entrepreneur-parole.目前没有哪个签证类型特别适合移民状态下的创业者。依赖于假释权利的 IER 是在立法途径都走不通的情况下提出的。立法措施可能更可取，但到目前为止其都被证明是政治上不可行的。

[23] *2010 年一份给国会的报告指出，有 24. 2 万张未使用的亲属移民绿卡最后被用于积压的职业移民申请，而国会通过特别立法重新发放了 18 万张绿卡，而留下 326,000 张绿卡被白白浪费了。《公民和移民局调查官：2010 年年报》，美国国土安全部（2010 年 6 月 30 日）https://www.dhs.gov/xlibrary/assets/cisomb_2010_annual_report_to_congress.pdf.今天这个数字可能更高，但 DHS 没有公布最新统计数据。*

[24]以前的国会行动实例包括在《2000 年 21 世纪法案》和《2005 年真实身份法案》中纳入美国竞争力条款。参见 Pub. L. 106 -313, 114 Stat. 1251, 1254（2000 年）和 Pub. L. No. 109 - 013, 119 Stat. 231, 322（2005 年）。

[25]具体为 8 U.S.C. § 1151(b)(1)。

[26]参见 William Kandel 所著《职业移民申请积压》，国会研究处，4-5（2020 年 3 月 26 日），https://fas.org/sgp/crs/homesec/R46291.pdf.

[27] *同上如 12。*

[28] EB-5 签证需要至少在美国投资 90 万美元。参见 William R. Kerr 所著《全球人才和美国移民政策：工作论文》，20-107，哈弗商学院，14（2020 年）），https://www.hbs.edu/faculty/Publication%20Files/20 -107_0967f1ab-1d23-4d54-b5a1- c884234d9b31.pdf.

[29] 83 Fed. Reg. 24415，《取消外籍创业者特殊工作许可项目》，美国国土安全部（2018 年 5 月 29 日），https://www.federalregister.gov/ documents/2018/05/29/2018-11348/removal-of-international-entrepreneur-parole-program.

[30]参见 Oren Etzioni 所著《特朗普行政命令在 AI 技术上的缺失：美国需要特别签证项目，旨在吸引更多 AI 专家》，《连线杂志》（2019 年 2 月 13 日），https://www. wired.com/story/what-trumps-executive-order-on-ai-is-missing/.

第十一章　加速 AI 创新

在 AI 创新中的领导地位

提升并调配联邦 AI 研发资金支持

开放全国性的 AI 研发基础设施拓展获取渠道

加强公私合营

应对人类的共同挑战

要维持美国在人工智能（AI）的世界领导地位，美国政府必须恢复其在国家层面上的创新投入优势。这需要在人工智能研发上进行大量新的投资，并建立全国性的人工智能基础研究设施，推动人工智能技术资源为大众所使用。国会议员必须达成协议，在今后数年内投入数百亿美元。这些投资的回报将改变美国的经济、社会和国家安全。

当我们的人工智能生态系统中存在阻碍创新的情况出现，结合目前中国采用国家导向人工智能项目的背景来看，此时缺乏国家层面的紧迫性是非常危险的。美国的人工智能技术开发是集中在少数地区的少数组织手里，研究途径有限。商业化议程影响着人工智能的未来，然而目前过于集中在一个方向：机器学习（ML）[1]。尽管美国政府采取了相应举措但相对于这个领域所具有的变革潜力，政府资金支持依然不足，这限制了政府朝着利于公益的方向引导人工智能的研究，同时制约了在一系列人工智能发展方向上取得进步的能力[2]。这也导致了人工智能的创新环境基础比较薄弱。

“当我们的人工智能生态系统中存在阻碍创新的情况出现，结合目前中国采用国家导向的人工智能项目的背景来看，此时缺乏国家层面的紧迫性是非常危险的。”

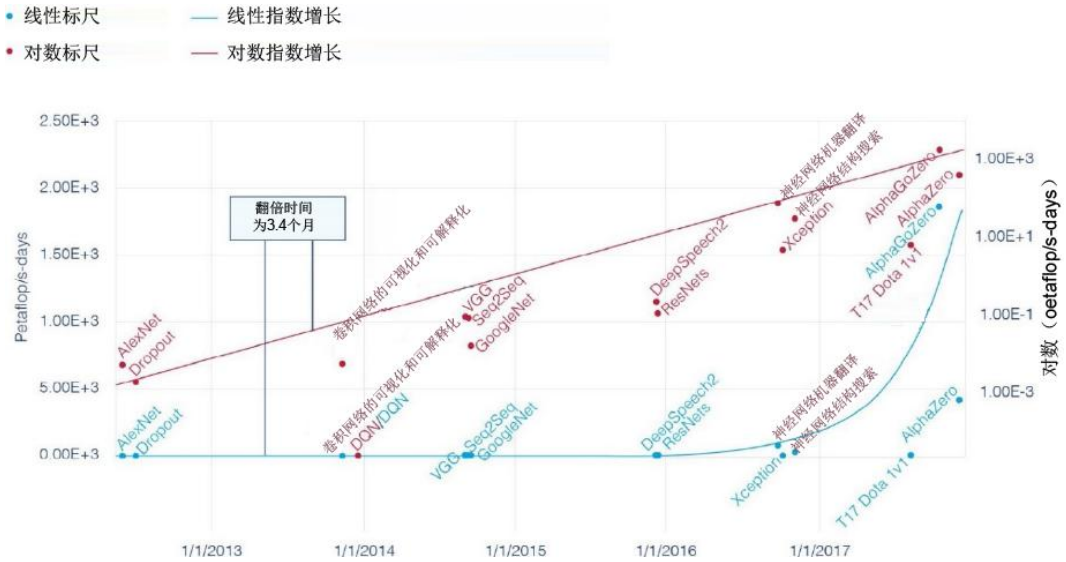
导致上述趋势的原因是资源。不断下降的云计算成本和的可获取的开源平台降低了核心机器学习的进入壁垒。然而，这同样也促成了复杂模型的实现，这些模型需要广泛的训练数据（通常存在于私人控制的数据集或知识图谱中）、巨大的算力和大量的软硬件工程[3]。如今，这些先决条件限定了人工智能研究的前沿技术范围，并限制了可在这个领域做出贡献并探索人工智能前沿技术的美国研究人员数量。

美国人工智能技术创新的关键在于独创性而非获取方式

人工智能行业的主流趋势在以下 5 个方面威胁着美国的技术竞争力：

- **人才流失。**人才从学术机构流失到私营部门，会让美国在基础人工智能研究上的优势根基，美国大学出现空心化的危机[4]。联邦资金支持的增长没有与该领域的增长保持同步，导致科研经费申请成功率较低，从而研究者们在编写计划书的文字上花费更多时间[5]。然而，学术专家和他们的学生不仅仅是因为官僚主义作风更少、财政激励举措更多才选择进入到大型科技公司。私营部门越来越成为开展前沿技术研究的最佳场所，他们拥有最佳的计算和数据资源。其结果就是削弱了下一代人工智能领导者在产业界和学术界的教育基础，让整个人工智能研究议程都变得艰难[6]。

最大规模深度学习模型训练所需的算力（2012-2017）



来源：OpenAI，人工智能和计算（2018 年 5 月 16 日），https://openai.com/blog/ai-and-compute/

“petaflop/s-day”是算力单位，由一天中按每秒钟可以进行 10 的 15 次方运算得到的算力消耗。

在此没有展现 GPT-3 和更多近期的模型，因为他们的值过大而无法显示在标尺上。

- **多元化。**人工智能中的“贫富”差距越来越大，让该领域中缺失多元化发展的现象更加严重[7]，限制了该领域（包括系统）通过集体维护的平衡能力。

- **研究焦点。**美国科技公司要为其股东负责，因此从逻辑上他们就不会在那些对公司没有商业价值或经济利益的国家安全重要领域进行投资，或在基础研究上冒险下注[8]。虽然以回报为中心的投资可以流向利于公益事业或政府工作的应用，但依然存在缺口。20 年前，机器学习和基础算法的地位正是如此，虽然表面上看不到商业前景，只能由联邦研究资金维持，直到算力和数据过剩才改变了产业发展方向[9]。近期有研究发现，今天所使用的算法中，有 82%是来自联邦资助的非盈利性组织和大学；与之相比，仅有 18%是来自私营企业[10]。

- **竞争。**开发前沿技术的机器学习模型不断增长的成本，以及大概率会被领先的科技公司收购的现象，意味着人工智能初创公司在美国的成长道路越走越窄[11]。缺乏竞争让产业的创新能力和人工智能研发的全球竞争力下降。

- **地区差异。**科技公司的地区性集群效应，例如硅谷效应，通过知识共享和国内竞争来驱动创新[12]。然而，人口统计学表明，这种效应会造成某些区域比其他区域受益更大[13]。2015 到 2017 年，美国创新区域的就业机会中，超过 90%是由仅仅 5 个沿海城市创造的[14]。这种地区差异使得技术进步所带来的红利集中在少数地区，而错过了全国其余地区的创新潜能。

联邦政府有责任去扭转以上趋势。提出战略方向，提供持续的资源，联邦政府既作为资金供应者，又是技术消费者[15]。政府必须打破标准科研基金的老旧窠臼。让技术创新在多个部门的努力开花结果，并且为实际应用方案的转换发挥最大价值。但以目前联邦机构和研究部门的行动结果来看，难以将有前景的技术概念从实验室推向实际应用。

被动的国家干预方式过于依赖私营部门来推动创新和导向研究方向，依赖商业创新“孵化”政府应用，采用被动方式是无法赢得这场战略竞争的，也无法充分利用人工智能技术的变革潜能。美国必须通过其在产业界和学术界合作中的政府领导地位，增加美国人工智能创新环境的多元化、竞争力和可达性。为此，可通过大量注入新的研发资金改变这一现状。

“美国必须通过其在产业界和学术界合作中的政府领导地位，增加美国人工智能创新环境的多元化、竞争力和可实现性。”

*提升联邦人工智能研发资金支持，做好调配工作。*为人工智能的研发进行大胆而全面的长期投资，将培育全国的人工智能创新精神，驱动突破创新的产生。在全面战略的指导下，注入持续资源支持，通过多元化分配机制，支持大批人工智能技术方法和人工智能在其他领域的新颖应用，进而实现由美国研究人员推动实际应用的目标。具体而言，美国应做到以下几点：

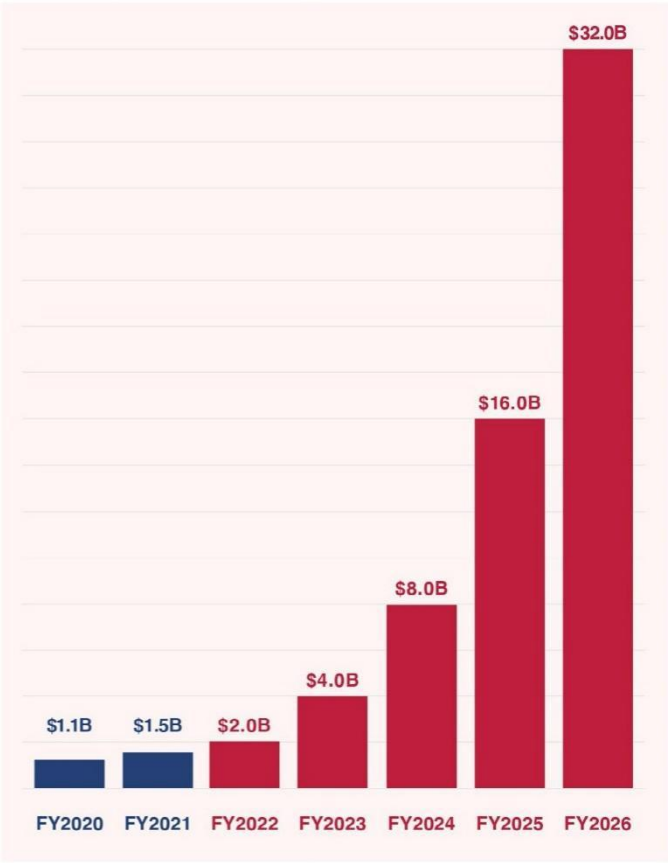
- **建立一个国家技术基金会（NTF）。**在取得成功的现有组织（例如 NSF 和 DARPA）基础上成立一个新的独立组织，采取补充措施，更激进地推动科技向工程方向发展。NTF 将在国家层面上推动技术进步，聚焦于在工艺成熟度的层面产生价值，优先关注由消费者需求而引发的概念，为试验和测试建设基础设施，支持研发成果的商业化。这需要一个组织可以承担较高风险，敢于在创新型观点和研究人员身上放手一搏。

- **增加用于非国防人工智能研发的联邦资金支持，按年复合增长率计算，每年翻倍，到 2026 财政年度达到每年 320 亿美元。**这个水平的人工智能支出相当于联邦政府在生物医学研究上的支出[16]。整体而言，政府应该至少支出 GDP 的 1%用于研发，以巩固科技领域的创新基础地位[17]。额外的资金支持应用于加强基础和应用研究，增设奖学金项目，支持用于研究的基础设施，并实现一些重点机构的覆盖：

  - 国家技术基金会（拟成立的）
  - 能源部
  - 国家科学基金会或国立卫生研究院
  - 国家标准与技术研究院
  - 国家航空航天局

人工智能研发投资水平

	FY2020	FY2021
NSF	$518.3M	$831.2M
NIH	$193.9M	$176.8M
DOE	$171.8M	$174.4M
USDA	$54.9M	$129.6M
DHS	$50.4M	$31.3M
FDA	$39.0M	$38.0M
NASA	$28.5M	$28.8M
NIST	$27.6M	$52.7M
DOT	$17.1M	$16.3M
VA	$14.1M	$14.1M
DOI	$5.9M	$4.2M
NIJ	$3.0M	$3.0M
NOAA	$1.6M	$1.6M
国库部	$0.6M	$0.6M
总计	$1.127B	$1.503B



DHS – 国土安全部
DOE – 能源部
DOI – 内政部
DOT – 交通部
FDA -食品药品监督管理局
NASA - 国家航空航天局
NIH - 国立卫生研究院
NIJ -国家司法研究所
NIST - 国家标准与技术研究院
NOAA -国家海洋与大气管理局
NSF -国家科学基金会
Treasury – 国库/金融犯罪执法网
USDA – 美国农业部
VA -退伍军人事务部

来源：网络和信息技术研发项目，2021 财政年度总统追加预算，国家科学和技术委员会（2020 年 8 月 14 日）https://www.nitrd.gov/ pubs/FY2021-NITRD-Supplement.pdf.

- **资金优先投入支持人工智能研发的关键领域。**增设的联邦资金应该优先投入对于未来支持国家安全和经济稳定的技术和至关重要的人工智能研发领域，并支持私营部门可能很少关注的领域。投资应由新成立的国家人工智能倡议办公室[18]统筹，以投资组合方式进行投资，聚焦在推进基础科学、解决特定问题，并推动商业化突破。

新颖的机器学习（ML）方向

当学习从一个任务或领域转移到另一个，以非监督式或半监督式方式对监督式机器学习采取非传统方式。

测试和评估、验证和确认（TEVV）

如何描绘一个 AI 系统的性能特征使之更好理解，包括在涵盖多系统的系统背景下和新环境中用于预测 AI 系统行为的改进方法。

强劲且有弹性的 AI

培养能够应对不利条件的方法，包括多种对抗性攻击，并改进可以评估脆弱性、豁免权和公平性的类型和水平的先进方法。

复杂的多个体场景

加强对 AI 系统交互群体的理解，包括对对手弱点和与对手缓和的研究，并在各种复杂场景中应用博弈理论。

集成 AI、建模、模拟和设计

推动利用模拟作为数据和场景来源进行 AI 系统培训和测试，并在科学发现和工程中将 AI 作为生成设计的引擎。

高级场景理解

进化感知模型，纳入多来源和多模式信息，在一系列复杂、动态的环境和场景中支持提高的可操作意识和洞察力。

保护个人隐私

通过可以实行匿名的技术和工艺的进步，例如多方联邦学习，在 AI 开发和应用中，确保数据使用的隐私保护。

AI 系统风险评估

发展支持风险评估的能力，包括评估可审核性、可解释性和可靠性程度的标准方法和指标。

更通用的人工智能方向

研究的挑战和人类智能的神秘之处，包括以非监督式方式进行学习的能力、累积和应用常识的能力、建立因果模型的能力、行使自我意识的能力以及概括知识的能力。

提高的人与 AI 互动和合作

在人与 AI 互补性、增加人类推理的方法、在混合系统中的流畅切换，以及 AI 技术对人类意图和通讯的理解方面取得进展。

AI 自动系统

发展系统在复杂的、不可预测的环境下独立完成任务，或人类操作员监督最小化时完成任务的能力。

- **国家人工智能研究机构的数量达到 3 倍。**政府应该在各地区和各研究领域将现有的联邦政府资助的国家人工智能研究机构的数量提升 3 倍[19]。这将增加学生和教员、国家实验室研究人员和非盈利研究组织的培训和研究机会。

- **投资于可以改变人工智能技术领域的人才。**与此同时，NSF 或拟成立的 NTF 应投资于顶级人工智能研究人员和跨领域团队，在那些可能带来无法预期突破的人和创造性理念上加大投入。

> “计算环境、数据和测试设施可为大众所使用，为产业巨头和精英大学之外的研究人员提供在推动人工智能前沿技术上发展的能力。”

建议

*开放全国性的人工智能基础设施使用渠道。*计算环境、数据和测试设施可为大众所使用，为产业巨头和精英大学之外的研究人员提供在推动人工智能前沿技术上发展的能力。支持人工智能领域更公平的成长，在全国范围传播人工智能相关知识，并在更广泛的领域中应用人工智能技术，夯实美国人工智能技术创新的基础。此类全国性基础设施应具有 5 个关键要素：

- **国家人工智能研究资源（NAIRR）**[20]。为减少“算力差异分化”[21]，NAIRR 将为研究人员和学生提供补贴，用于合理调配计算资源的使用。与实现了人工智能功能的政府和非政府数据集、教育工具和用户开展协作。NAIRR 应该利用联邦政府云平台按公私合营原则运行[22]。

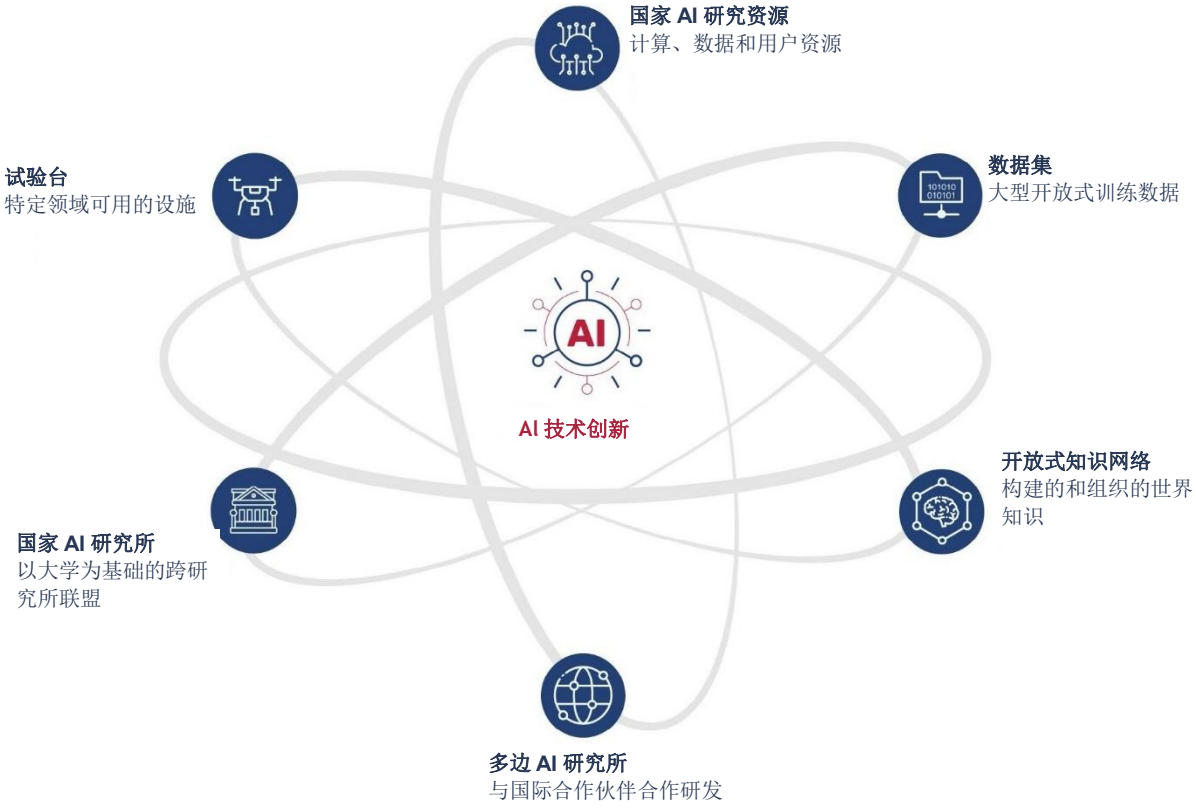
- **建设特定领域的人工智能研发试验平台。**在各类联邦机构资助下，这些试验平台将提供可适用的设施，建立基准标准，围绕符合公益的人工智能应用进行建设。

- **大规模开放式训练数据。**这应该包括复杂数据集的管理、托管和维护，对私营部门和学术界分享数据集的激励政策，对数据工程师和科学家对于开发政府所有的公众数据项目进行资助，供人工智能研究团队使用。

- **开放式知识网络。**由科学和技术政策办公室协调，此类资源将有助于构建和组织世界人工智能知识系统的开发，实现人工智能系统行之有效地运行[23]。

- **多边人工智能研究所。**与来自关键盟友和合作伙伴的研究人员进行合作研发（在本文第 15 章会做进一步阐述）。

这些资源互为补充，实现了数据、试验、测试和知识网络的良性循环，将人工智能技术创新和应用推向更广泛的挑战性问题和研究领域。



*利用公私合营双方的优势。*美国在如人工智能一类技术上的领先地位取决于更紧密的公私合作方式和关于美国全球竞争力的共同责任感。要促进创新，并加快具有全球竞争力的公司在战略新兴行业的增长，政府应采取以下措施：

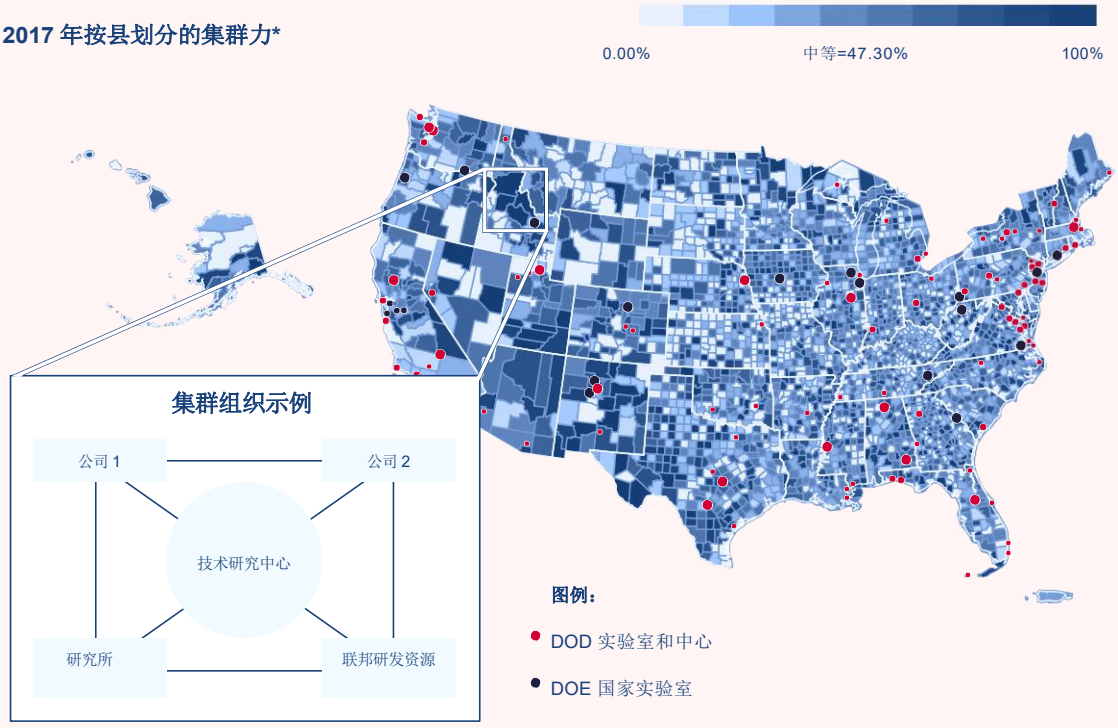
- **为人工智能和其他战略技术创造市场。**在各政府机构应用人工智能可以节省纳税人的钱，提升公共服务质量。有些应用可以直接从私营部门获取，有些对于政府

任务则需单独设计。通过加快人工智能技术在各联邦机构中的应用，政府可推动在人工智能应用中产生的额外商业投资，既有益于国家安全，也有益于公益[24]。

- **形成一个聚焦在战略新兴技术的地区性创新集群。**政府应为地区性创业集群指定发展方向，应聚焦于人工智能的战略新兴技术上，在对美国整体竞争力影响至关重要的行业中，培育小公司成长。通过竞标过程建立集群，创业集群将为产业界和学术界的工作者提供税收优惠、科研补贴和联邦研发资源。

地区性创新集群

2017 年按县划分的集群力*



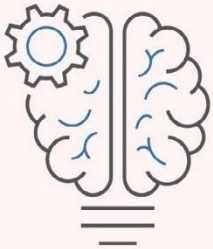
集群力是指在一个职业高度专业化地区的职业转换比例。这是选择地区性创新集群地点时，应考虑的众多重要因素之一。

图片来源：美国集群分布专题研究，战略和竞争力研究所，哈弗商学院。数据来源：美国统计局

私营部门应该：

- **资助人工智能竞争联合体。**私人公司应该建立一家非盈利组织，在 5 年内资助 10 亿美元，扩大人工智能研究机会，并支持人工智能技能培训和教育。这个以捐赠方式资助的组织应聚焦在通过人工智能研究和人工智能技能教育来促进经济机会。此类促进人工智能教育和创业精神的举措，将有助于缩短数字“贫富”差距。

应对人类的一些最大挑战。



应对人类的一些最大挑战。
通过聚焦在解决影响百万民众生活的现实人类困难，我们可以找到政府、学术界和产业界三方联盟新的存在理由，让公众对野心勃勃的 AI 研究继续予以支持，并延续美国在 AI 技术创新上的领导地位。可以改善社会福利和发展前沿科技的、前景光明的举措例证包括：

提高长期生活质量。
- 可以帮助老年人独立生活更长时间的 AI 技术，协助管理健康和日常任务，改善生活质量。

这包括在生物医药方面应用 AI 技术，解决急性、慢性病，提高老人健康。

教育革命和终身学习，用 AI 工具对难度适中的教育、培训和再培训进行个人化设置，直观评估学习进展，优化标准课程，改善个人学习成功效果。

改革能源管理。智慧城市基础设施，可对能源需求浪涌和紧急事件做出有效响应（包括人为和自然灾害）。

有效针对灾害做出预测、建模、备灾和响应。
- 准确、近实时天气、地震和消防线路检测及优化的预测，有助于应急响应援助和有限资源优化利用计划。
- 自动机器人，可在自然灾害或人为灾害后立即投入使用，进行搜寻、救援和清场，为现场急救人员和危险品专家提供多重支持。

研究太空前沿技术。自动 AI 的太空船、栖息地和设施，无论是否有人类干预，都可以确定并解决问题，实现更多更灵活的太空探索。

“通过聚焦在解决影响数百万民众生活的现实问题，我们可以找到政府、学术界和产业界三方联盟存在的新理由。”

第 11 章 – 尾注

[1]一篇关于 AI 的 arXiv 论文发现私营部门的基础 AI 研究主题比广泛的出版物资料库中的 AI 论文更为狭隘，主要集中在深度学习和支持深度学习的计算基础设施上。而且，研究发现与产业界合作更紧密的精英学术机构的研究主题集中度也与此类似地比较狭隘，导致美国 AI 研究环境偏离了多元化，而其他国家依然保持着多元化风格。参见 Joel Klinger 等人所著《AI 研究的狭隘化？》ArXiv （2020 年 11 月 18 日），https://arxiv.org/pdf/2009.10385.pdf.产业投资提升了硬件专业化程度，进一步让商用使用场景处于优先地位，导致追求主流以外的方法变得更加昂贵，参见 Sara Hooker 所著《硬件彩票》，ArXiv （2020 年 9 月 21 日），https://arxiv.org/ pdf/2009.06489.pdf.

[2]特朗普政府在 2021 财政年度提出的非国防 AI 研发预算为 15 亿美元，在 2020 年财政年度仅仅不到 10 亿美元的支出基础上有所增长。参加网络和信息技术研发项目，《2021 财政年度总统追加预算》，国家科学和技术委员会，4、12（2020 年 8 月 14 日） https://www.nitrd.gov/ pubs/FY2021-NITRD-Supplement.pdf. 2021 财政年度《国防授权法案》所包含的《2020 年国家 AI 倡议法案》构建了一个更具战略性方法的组织来管控 AI 技术，其做法是在科学和技术政策办公室内部建立一个国家 AI 倡议办公室以及相关顾问团和跨部门组织。参见 Pub. L. 116 -283, William M. (Mac) 2021 财政年度《国防授权法案》，134 Stat. 3388 (2021).

[3] OpenAI 估计，自 2012 年后，在最大型的 AI 训练中使用的算力值每 3.4 个月会翻番。参见 Dario Amodei 和 Danny Hernandez 所著，《AI 和算力》，OpenAI（2018 年 5 月 16 日） https://openai.com/blog/ai-and-compute/.基于 OpenAI 的预测，按照当前模型训练价格的增长速度来算，“在 4 年内，最大型模型的训练费用将超过火箭发射到轨道的费用。” 参见 Yaroslav Bulatov 所著，《大型 AI 和模型共享》，媒介（2019 年 7 月 20 日） https://yaroslavvb.medium.com/large-scale-ai-and-sharing-of-models-4622ba59ec18. AI 技术在机器人或真实世界中的应用开发，还需要额外资源，包括训练算法用的复杂建模和模拟能力，以及试验用的特殊设施。

[4]近期研究发现，2004 年到 2018 年之间，有 131 名教授从大学进入产业界，90 名教授则保持双重身份，在大学继续有兼职岗位。研究也证明了此类举动对这些大学的学生从事 AI 创业产生了不利影响。参见 Michael Gofman 和 Zhao Jin 所著 *《人工智能、教育和创业》*，SSRN，2（2020 年 10 月 26 日） https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3449440.商界的高薪水让研究人员脱离学术轨道：2019 年，北美有 57%的 AI/ML 博士毕业生进入产业界，其他的则在学术界继续攻读博士后、从事研究或作为教职员工。参见 Stuart Zweben 和 Betsy Bizot 所著《2019 年托比调查》，技术研究协会，11（2020 年 5 月） https://cra.org/wp-content/uploads/2020/05/2019 -Taulbee-Survey.pdf.

[5]例如，NSF 为计算科学提供了 85%的联邦资金支持，2019 年在核心 AI 研究中资助了 1.88 亿美元，但没有预算资助另外价值 1.78 亿美元的被高度评价的计划书。比起 2018 年算是进步，2018 年 NSF 资助了 1.65 亿美元，但还有 1.85 亿美元的被高度评价的工作未被资助。而且，NSF（与农业部合作）在 2020 年资助了 7 所国家 AI 研究所，但还有超过 30 家被认为值得支持的研究所未得到资助。参见《NSF 给 NSCAI 的陈述》（2020 年 1 月）。

[6]持有双重身份的计算科学教职员工在产业界花费的时间占比从 20%提高到 50~80%，这对他们的学术职责也产生了影响，包括招收学生、课程论文和研讨会工作，并且他们把产业界需求当作研究生的工作重点，地位高于重要的基础研究。参见 Shwetak Patel 等人所著《在技术研究中不断演变的学术界/产业界关系》，计算领域联合体，3（2019 年 6 月）https://cra.org/ccc/wp-content/uploads/ sites/2/2019/06/Evolving-AcademiaIndustry-Relations-in- Computing-Research.pdf.

[7]年度托比研究跟踪了计算科学（CS）领域，发现女性占 CS 学士学位毕业生的 21%，占 CS 博士毕业生的 20.3%，而国内没有被充分代表的少数裔占 CS 学士学位毕业生的 14.7%，占 CS 博士毕业生仅 3.1%。Stuart Zweben 和 Betsy Bizot 所著《2019 年托比调查》，技术研究协会，4、5、22（2020 年 5 月）https://cra.org/wp-content/uploads/2020/05/2019 -Taulbee-Survey. pdf.在 AI 研究领域的狭隘化趋势可能让这种情况雪上加霜。参见 Nur Ahmed 和 Muntasir Wahed 所著《AI 的去大众化：人工智能研究中的深度学习和算力差异分化》，ArXiv （2020 年 10 月 22 日）， https://arxiv.org/abs/2010.15581.

[8] See, e.g., Joel Klinger, et al., *A Narrowing of AI Research?*, ArXiv (Nov. 18, 2020), https://arxiv.org/pdf/2009.10385.pdf; Sara Hooker, *The Hardware Lottery*, ArXiv (Sept. 21, 2020), https://arxiv.org/pdf/2009.06489.pdf.

[9]在每个 AI 领域的开始阶段，联邦政府都会插手，呵护研究的成长。美国空军通过兰德公司资助 Herbert Simon

和 Allen Newell 的工作，这二者在 1956 年编写了第一代成功的 AI 计算机程序——逻辑理论家。参见 Mariana Mazzucato 所著《创业状况：公众和私营部门揭秘》，Anthem 出版社（2013 年）。国防高级研究计划局（DARPA，后来为 ARPA）资助了 Charles Rosen 的工作，Charles 在 1972 年开发了第一个自动巡航的机器人“Shakey”。参见《Shakey 机器人》，DARPA（最近登录时间为 2020 年 12 月 30 日），https://www.darpa.mil/about-us/timeline/shakey-the-robot.强化学习是如今商用应用方法的基础，当初是依靠 NSF 的支持才撑过了 1990 年代的“AI 寒冬”。参见 Andrew Barto 所著《NSCAI 员工与 NSF 的纠葛》（2019 年 8 月 8 日）。DARPA 对图像理解研究长达 30 年的资助奠定了如今自动驾驶能力的基础。参见 DARPAtv《DARPA 人工智能讨论会开场短片》，YouTube （2019 年 3 月 12 日）。https://www.youtube.com/watch?v=FTaW6ZJ9oyQ. 2000 年代中期，DARPA 运行的 PAL 项目引发了对第一代人工智能助手的开发，最终发展成为 Siri。参见 DARPAtv《DARPA 和 AI：有远见的先锋和拥护者》，YouTube （2018 年 12 月 7 日）。 https://www.youtube.com/watch?v=ri5gOjYgLns.

[10] Neil C. Thompson, et al., *Building the Algorithm Commons: Who Discovered the Algorithms that Underpin Computing in the Modern Enterprise?*, Global Strategy Journal at 4 (2020), https://onlinelibrary.wiley.com/doi/epdf/10.1002/gsj.1393.

[11]例如，非精英大学和 AI 初创公司难以承受复杂 ML 模型训练所需的计算资源和数据的高昂费用。参见 Nur Ahmed 和 Muntasir Wahed 所著《AI 的去大众化：人工智能研究中的深度学习和算力差异分化》，ArXiv（2020 年 10 月 22 日），https://arxiv.org/abs/2010.15581. 2013 年至 2018 年，90%的硅谷 AI 初创公司被大型科技公司收购。参见 Ryan Kottenstette 所著《硅谷公司正在弱化人工智能的影响力》，TechCrunch （2018 年 3 月 15 日）https://techcrunch.com/2018/03/15/silicon-valley-companies-are-undermining-the-impact-of-artificial-intelligence/. 在美国的专利清单上（除去采用专利），霸榜的同样是这些大型科技公司。参见 Al AuYeung，《谁在赢得 AI 竞赛？》，IPWatchdog（2020 年 2 月 1 日） https://www.ipwatchdog. com/2020/02/01/winning-ai-race/id=118431/.

[12] Michael Porter, *Clusters and the New Economics of Competition*, Harvard Business Review (Nov.- Dec. 1998), https://hbr.org/1998/11/clusters-and-the-new-economics-of-competition.

[13] William R. Kerr & Frederic Robert-Nicoud, *Tech Clusters*, Journal of Economic Perspectives at 63 (Summer 2020), https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.34.3.50.

[14] Specifically, Seattle, Boston, San Francisco, San Diego, and San Jose. Robert D. Atkinson, et al., *The Case for Growth Centers: How to Spread Tech Innovation Across America*, Brookings (Dec. 9, 2019), https://www.brookings.edu/research/growth-centers-how-to-spread-tech-innovation-across- america/.

[15] NSF 和其他政府机构所做的工作令人钦佩，NSF 利用其资源，鼓励多元化研究，为 AI 创新创造规模经济，但就目前的努力而言，他们尚无法形成战略效应，这是联邦的研发投资整体在下降的背景所决定的。近期联邦采取的其他突出举措包括 DARPA 的人工智能探索项目，这个项目可以在 18 个月时间内快速提供高达 100 万美元的资金，用于探索新 AI 概念的可行性。NSF 的国家 AI 研究所项目，在 2020 年资助了 7 家以大学为基础的跨研究所联盟，每年 400 万美元，连续 5 年，并计划在 2021 年发起对另外 8 家的资助。参见《NSF 的人工智能》，NSF（2020 年 8 月 26 日），https://www.nsf.gov/cise/ai.jsp.

[16]联邦对国立卫生研究院（NIH）的资金支持，从 2010 年的 300 亿美元增长到 2020 年的 410 亿美元。参见《NIH 预算历史：NIH 预算机制详情》，NIH 数据手册（2019 年 10 月）https://report.nih.gov/nihdatabook/category/1; *Budget*, NIH (June 29, 2020), https://www.nih.gov/about-nih/ what-we-do/budget.

[17] 1953 年，美国 GDP 的 0.72%用于研发。1957 年，当在苏联发射了斯普特尼克人造卫星之后，这个数字增长到 1.3%。1964 年研发支出达到了顶峰的 1.86%。而到 2017 年，又下降到低于 1953 年水平的 0.61%。参见《联邦研发预算仪表盘》，美国科技进步协会（最近登录时间：2021 年 1 月 14 日）https://www.aaas.org/programs/r-d-budget-and-policy/ federal- rd - budget- dashboard.

[18] 2021 财政年度《国防授权法案》所包含的《2020 年国家 AI 倡议法案》构建了一个更具战略性方法的组织来管控 AI 技术，其做法是在科学和技术政策办公室内部建立一个国家 AI 倡议办公室以及相关顾问团和跨部门组织。参见 Pub. L. 116 -283, William M. (Mac) 2021 财政年度《国防授权法案》， 134 Stat. 3388 (2021)。

[19] NSF 为 2020 年第一批国家 AI 研究所发放资金，资助 7 所以大学为基础围绕 AI 研究的基础和应用领域建立的跨研究所联盟，并计划在 2021 年资助第二批研究所，不仅支持跨部门的合作伙伴，也会支持私营部门的利益攸关人，并再发起对另外 8 家研究所的资助。参见《NSF 的人工智能》，NSF（2020 年 8 月 26 日），https://www.nsf.gov/cise/ai.jsp.

[20]按照 NSCAI 在美国《一季度建议》中的内容，国会在 2021 财政年度《国防授权法案》中采取了建立 NAIRR 的第一步，创建了一个工作组，为将来的 NAIRR 制定线路图。工作的成果将在任命工作组成员后 18 个月向国会汇报。参见 Pub. L. 116 -283, William M. (Mac) 2021 财政年度《国防授权法案》，134 Stat. 3388 (2021)，也可参见《一季度建议》，NSCAI，12（2020 年 3 月）. https://www.nscai.gov/previous-reports/.

[21]自从 2012 年深度学习爆发后，同时伴随着 AI 计算专用硬件的使用增长，被称为“算力差异分化”的现象随之出现，这是在 AI 前沿技术研究中所必须拥有的资源使用权上，出现的大型科技公司和精英大学与中小层级大学之间的不平等现象。参见 Nur Ahmed 和 Muntasir Wahed 所著《AI 的去大众化：人工智能研究中的深度学习和算力差异分化》，ArXiv（2020 年 10 月 22 日）， https://arxiv.org/abs/2010.15581.

[22]这种方式可以建立在成功的模型之上，例如新冠病毒高性能计算联合体 (https://covid19 -hpc-consortium.org/)和 NSF’s 的 CloudBank，(https://www. cloudbank.org/).

[23]这将建立在网络和信息技术研发（NITRD）项目大数据跨部门工作组之前承担的工作之上。参见《开放式知识网络：大数据跨部门工作组研讨会》，国家科学和技术委员会（2018 年 11 月） https://www.nitrd.gov/pubs/Open-Knowledge-Network-Workshop-Report-2018.pdf.这也是基于 NSF 对开放式知识网络进行聚合加速器追踪的持续努力之上。参见《NSF 聚合加速器奖励让工人同时受益于科学家、企业和非盈利组织》，NSF（2019 年 9 月 20 日）https://www.nsf.gov/news/special_reports/announcements/091019.jsp.

[24]国会在《2021 年度综合拨款法案》中迈出了重要一步，要求联邦总务署启动一个五年计划，也称为“AI 卓越中心”（AI CoE）,在其他任务中“推动人工智能技术在联邦政府中的采用”。AI CoE 有助于将各联邦机构的零散努力连结起来，形成可观的政府特定的 AI 应用市场。参见规则委员会打印版 *116 – 68，《对 H.R. 133 参议院修正的内务修正案文本》，美国内务规则委员会* 378-81 *（2020 年 12 月 11 日）* https://rules.house.gov/sites/democrats.rules. house.gov/files/BILLS-116HR133SA-RCP-116 - 68.pdf （具体参阅《2021 年度综合拨款法案》的第 103 节）。此外，国防创新小组（DIU）在 AI 和其他战略技术交汇处，通过其基于项目的方式为创造市场发挥了作用。AI 应用的关注领域包括太空系统、高级诊断、半导体/高级硬件以及在本文第 16 章中由 NSCAI 确认的其他关键技术。 DIU 的经验是，为战略技术制造市场，可以从国防部（DoD）和其他政府机构开始，采取的方式为：（a）合同灵活；（b）与公司的技术发展规划一致；及（c）通过规模生产的机会形成财政激励。参见《五年间 DIU 产生的变革性影响》，国防部（2020 年 8 月 27 日） https://www.defense.gov/Explore/News/Article/Article/2327021/diu- making-transformative-impact-five-years-in/.

第十二章　知识产权


产权问题签发行政命令知识产权政策是国家安全的重中之重
就 AI 和新兴技术的知识
制定计划，改革和建立知识产权政策和制度
评估“知识产权考量事项”
提出行政和立法行动建议
纳入国家安全、经济和技术竞争力战略之中

中国正在借用和利用知识产权（IP）政策，将其作为国家新兴技术战略中的关键工具。在保障国家安全、经济利益和技术竞争力问题上，美国未能以相同的高度认识到知识产权的重要性。美国没有建立全面的知识产权政策，以激励对人工智能（AI）和其他新兴技术领域进行投资[1]，并保护这些技术的创生[2]。美国的知识产权政策存在诸多失位，其一是当前美国专利资格和专利性原则导致的法律不确定性，其二是缺乏对中国围绕其知识产权机构构建的国内和地缘政治战略的有效回应[3]，其三是缺乏有效的数据保护政策。以上都可能使美国失去其在知识产权的全球领导地位。与此同时，中国通过加强其知识产权制度[4]，时刻准备“填补”因为美国知识产权保护弱化而造成的“空缺”，特别是在专利上，因为美国已经失去了其“在新技术创新上获得稳定而有效的知识产权的相对优势”[5]。这种明显的政策不对称对美国国内和国际都有多重重大影响。。

首先，美国法院严格限制了哪些类型的计算机实现和生物技术相关的发明可以根据美国专利法得到保护[6]。自 2010 年以来，关键的人工智能和与生物技术相关的发明就失去了专利保护。面对获得和保留专利保护的不确定性时，发明者会转而寻求商业秘密保护。但是商业秘密不易推动创新市场，与专利不同，其在公共知识领域中的贡献有限[7]。虽然这些影响可能不会即刻发生，但对于人工智能和其他新兴技术的发展和竞争力的长期影响依旧令人担忧[8]。

其次，中国已经实现了其增加专利申请和专利颁发数量的战略政策目标，打造了“赢得”创新竞赛的故事。2019 年，在中国国家知识产权管理局（CNIPA）登记的“创新”专利申请总数量，大约是在美国专利商标局（USPTO）登记的实用新型专利申请数量的 3 倍[9]。在世界知识产权组织（WIPO）的国际专利合作条约（PCT）框架下，中国的国际专利申请数量也处于世界领先地位[10]。至关重要的是，中国如今经常被当作是人工智能创新发明国内专利申请登记数量的领先者[11]。在全球范围内，源自中国的人工智能专利申请数量已经超出了源自美国的数量，尤其是在近些年中[12]。

“美国没有建立全面的知识产权政策，以激励对人工智能（AI）和其他新兴技术领域进行投资，并保护这些技术。政策存在诸多失位...其结果就是：美国可能会丢掉其在知识产权的全球领导地位。”

产权趋势：

中国对于 AI 和新兴技术的国家知识产权制度：

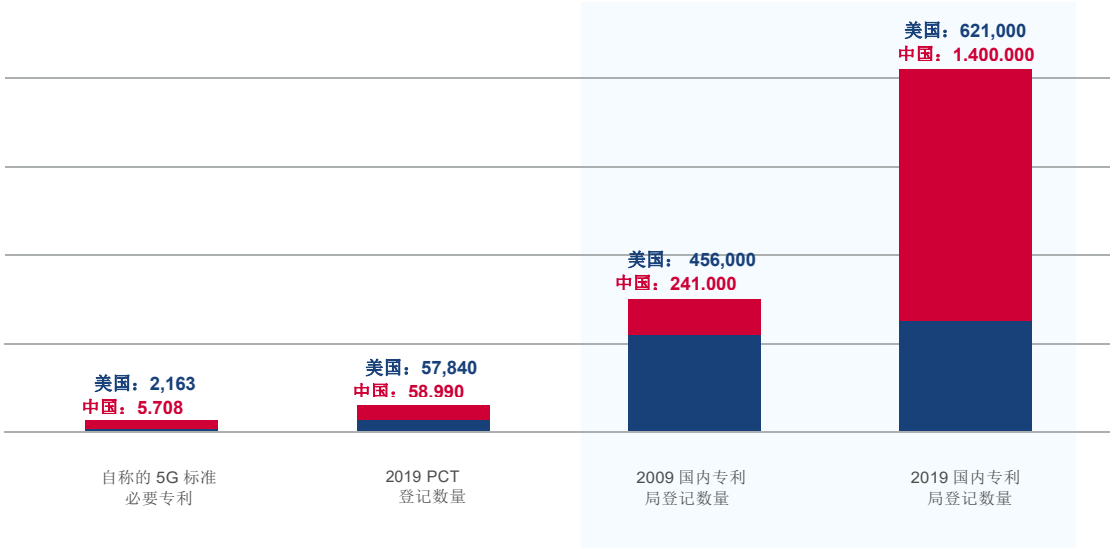
中国战略新兴产业发展的十三五规划中，清晰阐述了与知识产权相关的新兴技术目标：

- 修订专利法和版权法
- 通过快速维权中心加强知识产权保护
- 制定新兴技术知识产权的战略发展规划
- 改善海外知识产权，并支持公司参与海外并购

专利登记的激励因素包括：

- 专利补贴
- 颁发专利奖励
- 省市级政府设定战略配额
- 在政府采购过程中，具有中国知识产权的公司享有优惠政策

对专利侵权发布初步禁令、加大惩罚性赔偿（对恶意侵权者给予 5 倍赔偿惩罚）、设立具有有效解决方案和低诉讼费的专门知识产权法庭来提高专利保护



《“十三五”国家战略性新兴产业发展规划》的 CSET 翻译版，中国共产党中央委员会和中华人民共和国国务院（2016 年 11 月 29 日出版）（CSET2019 年 12 月 9 日翻译）https://cset.georgetown.edu/research/national-13th- five-year-plan-for-the-development-of-strategic-emerging-industries//；参见 Eric Warner《中国的专利申请和创新：激励措施、政策和成果》，兰德公司，17-18（2014 年 11 月），https://apps.dtic.mil/dtic/tr/fulltext/u2/a619128.pdf；《中国的商标和专利：非市场因素对专利登记趋势和知识产权系统的影响》，美国专利商标局（2021 年 1 月），https://www.uspto.gov/ sites/default/files/documents/USPTO-TrademarkPatentsInChina.pdf；参见 Ryan Davis《中国修订的专利法必知之四件事》，Law360（2020 年 11 月 5 日），https://www.law360.com/articles/1326419/；参见陶凯元法官《中国承诺加强知识产权司法保护，创造知识产权美好未来》，WIPO 杂志（2019 年 6 月），https://www.wipo.int/wipo_magazine/en/2019/03/article_0004.html.

注释：自主申报的 5G 标准必要专利数量是 2020 年 2 月的数据，代表了两个国家中最大的两家公司专利数总和。如美国的 2163 件为高通的 1293 件专利申请登记加英特尔的 870 件。而中国的 5708 件为华为的 3147 件和中兴通讯的 2561 件。这个数字也是指登记的标准必要专利，而非颁发的专利数量。参见 Jed John Ikoba 的《截止 2020 年 2 月，华为具有全球最多的 5G 专利》报告，Gizmochina（2020 年 6 月 2 日），https://www.gizmochina.com/2020/06/02/huawei-has-the-most-5g-standard-essential-patents-globally/ ；《2019 年，在 WIPO 知识产权服务强劲增长的背景下，中国成为最大国际专利申请国》，《条约和经济》，WIPO 媒体中心（2020 年 4 月 7 日），https://www.wipo.int/pressroom/en/articles/2020/ article_0005.html；至于中国国内专利

登记数量，按照中国国家知识产权局（CNIPA）的数字，“2009 年到 2019 年，它收到的创新型专利申请数量增长了 500%，从 24.1 万件增长到 140 万件（有趣的是 2018 年到 2019 年下降了 9%）。与之对比，同时期内美国专利和商标局的专利申请数量仅仅增加了 35%（从 45.6 万增长到 62.1 万件）。因此，2009 年美国专利的申请数量几乎是以 2 比 1 的比例远超中国，然而到 2019 年，这个比例完全颠倒了。中国专利申请数量的增长大部分是来自国内申请人的申请。2019 年 CNIPA 的 140 万件申请中，国内的申请几乎占了 90%（而美国专利和商标局是 48%）。”参见 Patrick Thomas 和 Dewey Murdick，《专利和人工智能：入门》，安全和新兴技术中心，10（2020 年 9 月）https://cset.georgetown.edu/wp-content/uploads/CSET-Patents-and-Artificial-Intelligence.pdf. 2019 年，中国的实用新型专利申请几乎达到 200 万件。参见同上，17。

“中国已经实现了其增加专利申请和专利授予数量的战略政策目标，打造了“赢得”创新竞赛的故事。”

第三，若忽略质量问题[13]，中国大量的专利申请文件可能会进一步伤害美国创新者，因为他们创造了大量的 “现有技术”（prior art，专利法中对全球科学技术知识的术语，指对一项发明进行评估以确定其是否为新发明），这大幅增加了检查专利申请中审查现有技术的数量，从而导致专利审查过程变得越来越难。与此同时，美国发明者也会发现获得专利比以前更难了，因为他们必须表明其发明没有在世界任何地方的现有技术刊物上公布过，包括在中国和国际上登记的中文专利申请[14]。当中国专利逐步主宰全世界专利局的现有技术搜索时，美国专利在全球范围现有技术检索中的主导地位将被蚕食[15]。

第四，与中国大量的专利申请登记一致，中国公司也在标准开发组织中将太多专利认定为“标准必要”专利，宣称这些专利是在实施技术标准时必须使用的专利[16]。虽然标准开发组织要求专利持有人对专利在未来标准中可能被认定为必要的专利进行自我确认，但这些组织还是把最终必要性的决定权留给私人企业进行许可谈判；若出现争议，则决定权归为法院[17]。这种“过度申报”标准必要专利（SEPs）的做法，进一步打造了中国在全球“赢得”诸如 5G 的标准化技术竞赛的故事，促使其他国家在自己的通讯基础设施中采取了中国的技术[18]。一个令人不安的结果就是美国公司必须向中国公司支付数十亿美元的专利使用费，否则将因为他们恶意侵犯中国公司专利权而面临着索赔及由此产生的诉讼[19]。

第五，对数据缺乏明确的法律保护或对数据所有权缺乏明确政策会阻碍创新和合作。特别是随着技术的发展[20]，数据保护制度的缺失会阻碍各方对开发数据集进行必要投资的积极性，而数据集对机器学习（ML）和人工智能系统至关重要[21]。此外，数据管理政策或所有权规则缺乏对知识产权类保护（例如规定最佳做法）也会削弱公司加入公私合营伙伴关系的意愿，而公私合营对于前沿技术的创新非常关键[22]。如果数据权利或所有权声明出现疑问，会让美国及其盟友与其他合作伙伴在人工智能研发上的合作出现问题[23]。

知识产权政策失位

美国政府需要解决因为全面知识产权政策缺失而导致的这些漏洞。如今，美国政府无法有效利用知识产权政策作为国家战略工具，以支持国家安全、经济利益、人工智能技术和新兴技术的技术竞争力。美国政府在知识产权政策的协调工作主要集中在知识产权执行和防止知识产权窃取方面。然而，美国缺乏一个机构或跨机构的部门，被授权同时制定和执行国家知识产权政策，以支持国家战略，并在国家战略中纳入知识产权政策。因此，美国缺乏统一的、立法授权的人工智能和新兴技术知识产权政策。这些政策将纳入到美国国家战略框架中，并应对来自其他国家例如中国的全球竞争。

必须将美国的知识产权法律和机构视为保卫国家安全利益、推动经济繁荣和技术竞争力的关键部分。美国至少应做到以激励、拓展和保护人工智能和新兴技术为目标，在国内外阐明自己国家的知识产权改革措施和政策，并完善相关措施和政策的制定。此类政策应由行政部门制定并提出，在这个过程中应该将对促进美国技术创新有重要作用的各个部门和机构整合起来。行政部门应该：

“必须将美国的知识产权法律和机构视为保卫国家安全利益、推动经济繁荣和技术竞争力的关键部分。”

*制定和执行国家知识产权政策，以激励、拓展和保护人工智能和新兴技术。*总统应该签署一份行政命令，确认知识产权作为一项国家重点事项，并制定一份全面计划，完善改革，制定知识产权政策和制度，以进一步推进国家安全、经济利益和技术竞争力战略。委员会建议该行政命令应指示副总统作为技术竞争力委员会（TCC）主席，或作为跨机构特别工作组组长，对这项工作进行监督。行政命令也应指示委员会秘书，必要时，在负责知识产权的商务部副部长和美国专利和商标局局长的配合下[24]，准备好改革和制定新的知识产权政策和制度的建议书，以激励、拓展和保护人工智能和新兴技术。该规划应包括知识产权政策变更的行政和立法行动提议，以实现这些目标，且应附有一份非详尽的“知识产权考量事项”清单的评估[25]。行政命令应指示副总统对商务部部长提供的知识产权政策、制度和改革建议书进行评估，确定哪些可以整合进国家安全、经济和技术竞争力战略中，并授权商务部部长推动这些提议的执行。

国家知识产权考量事项

- 专利资格性
- 对中国基于专利申请登记数量而打造的“赢得”技术竞争的故事进行反击
- 中国专利申请对美国专利和商标局和美国发明者的影响
- AI 公私合营和国际合作的障碍
- 数据的知识产权保护
- 与窃取知识产权的行为做斗争
- AI 的发明人身份
- 全球知识产权结盟工作
- 创新和知识产权生态系统大众化
- “标准必要”专利过程

第 12 章——尾注

[1]新兴技术进步需要大量投资。这些投资部分是公营的，但也需要广泛的私营投资。

[2] 对国家安全利益关键的技术包括 AI 和生物技术。本报告第 16 章中列出了 NSCAI 所提出的一份初步清单，包括对美国国家竞争力比较关键的新兴技术。

[3]《“十三五”国家战略性新兴产业发展规划》的 CSET 翻译版，中国共产党中央委员会和中华人民共和国国务院（2016 年 11 月 29 日出版）（CSET2019 年 12 月 9 日翻译） https://cset.georgetown.edu/research/national-13th-five-year-plan-for-the-development-of-strategic-emerging-industries/. 中国在推动实现创新和产业竞争力目标的过程中，继续对其知识产权体制进行广泛改革。参见 Mark Cohen《最近专利立法的 IPO 评论：解开一张错综复杂的网》，中国知识产权日志（2020 年 12 月 15 日） https://chinaipr.com/2020/12/15/ipos-comments-on-recent-patent-legislation- untangling-a-complex-web/.

[4]中国采取的行动包括确保 AI 和相关技术可以被专利所保护，提高专利侵权的罚判金额，继续为有效专利的侵权发布初步禁令，并设立具有有效解决方案和低诉讼费的专门知识产权法庭。参见 Kevin Madigan 和 Adam Mossoff《点金成石：专利资格原则是如何削弱美国在创新中的领导地位》，乔治梅森法律评论，943-946（2017 年 4 月 13 日），https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2943431（此后称《点金成石》），参见 Ryan Davis 中国修订的专利法必知之四件事》，Law360（2020 年 11 月 5 日），https://www.law360.com/articles/1326419/; 参见 Liaoteng Wang 等人《专利主题适任性标准对比：中国 VS 美国》，IP Watchdog（2020 年 6 月 12 日），https://www.ipwatchdog.com/2020/06/12/comparative-look-patent-subject-matter- eligibility-standards-china-versus-united-states/id=122339/；参见 Erick Robinson《中国新初步禁令规则之你需要知道的一切》，IAM（2018 年 12 月 21 日），https://www.iam-media.com/designs/everything-you-need-know-about-chinas-new-preliminary-injunction-rules; 参见陶凯元法官《中国承诺加强知识产权司法保护，创造知识产权美好未来》，WIPO 杂志（2019 年 6 月），https://www.wipo.int/wipo_magazine/en/2019/03/article_0004.html.

[5] 参见《点金成石》，955。

[6]参见《点金成石》。2019 年 1 月，美国专利商标局（USPTO）公布了初步专利资格指南，适用于在美国专利和商标局登记的专利申请审查，有争议地减少了因为专利申请审查和批准过程因为判定专利资格带来的不确定性。然而，美国联邦巡回法院的上诉法庭，作为对专利法案具有上诉管辖权的上诉受理法庭，认为其不受该指南的约束。参见 *Cleveland Clinic Found. v. True Health Diagnostics LLC*, 760 F 案，约 1013、1020 (Fed. Cir. 2019)（无先例）；*In re Rudy*, 956 F.3d 1379, 1383 (Fed. Cir. 2020)（有前例）（引用 *Cleveland Clinic Found.*, 760 F 案，约在 1021）。

[7]参见 *Crash Course《专利：专利是什么，为何专利有用》*， Ius mentis（最后登录时间 2020 年 12 月 30 日） https://www.iusmentis.com/patents/crashcourse/whatis/（因为专利公开发布发明细节，其他发明者可获得该发明的许可或考虑改进发明或围绕披露内容进行设计）；参见 Steven Hoffman & Calla Simeone《商业秘密保护和新冠病毒治疗：对联邦政策决定的观察及其对生物医药技术进步的潜在影响 》，JDSupra（2020 年 9 月 15 日） https://www.jdsupra.com/legalnews/trade-secret-protection-the-covid-19 -37383/（讨论对于在生物医药技术进步上使用商业秘密保护时，专利资格的不确定性带来的意义）。

[8]调查和产业报告证明投资已经偏离了专利密集行业。参见 Mark F. Schultz 《有效和可靠的专利系统对在关键技术中进行投资的重要性》，美国初创公司联盟和职业投资，24-37（2020 年 7 月） https://static1.squarespace.com/static/5746149f86db43995675b6bb/t/5f2829980ddf0c536e7132a4/1596467617939/USIJ+Full+Report_Final_2020.pdf.

[9]参见 Patrick Thomas 和 Dewey Murdick《专利和人工智能：入门》，安全和新兴技术中心，10（2020 年 9 月） https://cset.georgetown.edu/wp-content/uploads/CSET-Patents-and-Artificial-Intelligence.pdf（此后称《CSET 入门》）；美国日历年 1963-2019 专利统计图表，美国专利和商标局（2020 年 4 月） https://www.uspto.gov/web/offices/ac/ido/oeip/taf/ us_stat.htm.

[10]参见《CSET 入门》，11；参见 Aaron Wininger《中国超越美国成为 2019 年 PCT 国际专利的最大申请国》，国家法律评论（2020 年 4 月 7 日） https://www. natlawreview.com/article/china-surpasses-us-to-become-top-filer-pct-international-patent-applications -2 019.中国正在继续成为 2020 年的最大 PCT 申请国。参见 Aaron Wininger《中国 2020 年上半年专利数据显示中国可能继续为最大的国际专利申请国》，国家法律评论（2020 年 7 月 11 日）https://www.natlawreview.com/article/china-2020 -h1-patent-data-i ndicates-china-likely-to-remain-top-international-filer.

[11] AI 发明者，RS（最近登录日期：2020 年 12 月 30 日） https://uk.rs-online.com/web/generalDisplay. html?id=did-you-know/ai-innovators；参见 George Leopold《中国主宰了 AI 专利申请》，Enterprise AI（2020 年 8 月 31 日），https://www.enterpriseai.news/2020/08/31/china-dominates-ai-patent-filings/； 《CSET 入门》。

[12] 《CSET 入门》，9、12、n. 23。

[13]《中国的商标和专利：非市场因素对专利申请趋势和知识产权系统的影响》，美国专利和商标局，1（2021 年 1 月） https://www.uspto.gov/sites/default/files/documents/USPTO- TrademarkPatentsInChina.pdf; 参见 Jonathan Putnam 等人《中国的创新成果》，SSRN，32（2020 年 8 月）（有待修订），https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3760816.

[14] 参见 Jeanne Suchodolski 等人《创新战争》，北卡罗来纳州法律和技术杂志，201（2020 年 12 月），https://ncjolt.org/articles/volume-22/volume-22-issue-2/innovation-warfare/ （此后称《创新战争》）

[15] 参见 Rob Sterne《中国如何能根本性改变全球知识产权系统》IP Watchdog（2019 年 7 月 24 日），https://www.ipwatchdog.com/2019/07/24/china-changing-global-ip-system/id=111613/.

[16]过度申请已经出现在 5G 领域。参见 Matthew Noble 等人《决定哪家公司在领跑 5G 技术》，IAM（2019 年 7-8 月） https://www.twobirds.com/~/ media/pdfs/news/articles/2019/determining-which-companies-are-leading-the-5g-race. pdf?la=en&hash=8ABA5A7173EEE8FFA612E070C0EA4B4F53CC50DE;《面对中国的挑战：美国的技术竞争新战略》，中美关系科学和技术工作组，27、29（2020 年 11 月 16 日） https://china.ucsd.edu/_files/meeting-the-china- challenge_2020_report.pdf （此后称《面对中国的挑战》）。

[17] IEEE SA 标准董事会章程，IEEE 标准协会（最后登录时间：2020 年 1 月 15 日） https://standards.ieee.org/about/policies/bylaws/sect6 -7.html#loa.

[18]《5G 技术领导地位》，哈德逊研究所（2020 年 12 月 5 日）https://www.hudson.org/ research/16547-5 -g-technological-leadership;《创新战争》，201、n.130（中国企业认识到标准制定行动的重要性，以及参与论坛可以为使用和影响发展中的技术提供了合法手段）。

[19]因为标准必要专利（SEPs）可能涉及到数十万技术，授权费具有显著经济影响。参见《5G 技术领导地位》，哈德逊研究所（2020 年 12 月 5 日），https://www.hudson.org/research/16547-5 -g-technological- leadership （“如果法官或监管部门简单计算专利总数决定专利组合价值的度量标准，则专利计算可能对在美国创新经济中工作的企业产生不利影响。不考虑专利质量的差异，将有可能对一些价值较低的技术专利持有者过度补偿，而对具有突破创新的专利持有人则补偿不足”。）参见 Andrei Iancu（美国专利和商标局局长）在 2020 年秋季知识产权会议中心的发言（2020 年 10 月 7 日） https://cpip.gmu.edu/2020/10/20/cpip-2020 -fall-conference-day-one-recap/；参见 Muzammil Hassan 等《谁拥有核心 5G 专利？5G 标准必要专利的详细分析》GreyB（2020），https:// www.greyb.com/5g-patents/#The-State-of-Declared-5G-Patents；参见 Cody M. Akins《标准必要专利的过度申请》，德克萨斯法律评论（2020 年），https://texaslawreview.org/wp-content/uploads/2020/02/Akins.Printer.pdf.

[20] 参见 Mitchell Smith《美国和欧盟数据库的法律保护措施：给科技研究带来的含义》，SSRN（2010 年 5 月 23 日），https://papers.ssrn.com/sol3/papers. cfm?abstract_id=1613451；参见 Daniel J. Gervais《探索大数据和知识产权法之间的分界》，知识产权、信息技术和电子商务法杂志（2019 年），https://scholarship.law.vanderbilt.edu/faculty-publications/1095.

[21]在美国专利和商标局调查利益攸关人对 AI 技术知识产权观点的一份报告中，“评论家们基本平分为两个观点：新的知识产权可以解决 AI 创新问题，和当前美国知识产权框架不足以解决 AI 创新问题。然而，通常来说，没有看到知识产权需要变革的评论家们建议对 AI 技术发展进行监控，确保对 AI 技术的需求与 AI 技术发展同步。大部分观点要求新的知识产权聚焦在保护 AI 相关的数据上，特别是 ML。”美国专利和商标局关于人工智能和知识产权政策的公开观点，15（2020 年 10 月）https://www.uspto.gov/sites/default/files/ documents/USPTO_ AI-Report_ 2020 -10 - 07.pdf.

[22] 参见 Thomas E. Ayers《改变我们购买武器的方式将利于产业发展》，政府和纳税人，国防新闻（2019 年 11 月 20 日），https://www.defensenews.com/opinion/commentary/2019/11/20/changing-how-we-buy-weapons-will-benefit-industry-government-and- taxpayers/（讨论空军和供应商之间在知识产权保护问题上的紧张局面）

[23]也参见第 15 章的《行动的蓝图》。

[24]这些工作上，其他行政部门和机构以及美国版权局应资助并支持商务部部长。

[25]知识产权考量事宜的非详尽清单应该包括专利资格原则，对中国利用专利申请数量打造“赢得”AI 创新的故事进行反击，中国专利申请对美国专利和商标局专利审查过程和对美国发明者的影响，知识产权合约系统对公私合营和国际合作的阻碍，数据的知识产权保护，和窃取知识产权行为斗争，AI 的发明者身份问题，全球知识产权的协调，创新和知识产权生态系统和标准必要专利过程。

第十三章　微电子

重获微电子领域领先地位

保持微电子领域领先两代水平

境内尖端制造的多重来源

国家微电子战略

倍增的微电子研发资金支持

给予美国设备生产商减免税优惠政策

取得领先的微电子技术对于美国在人工智能技术上（AI）的领导地位至关重要。以下这些评估支持了这个观点：

- 硬件是人工智能堆栈和数据、算法及人才的基础要素[1]。
- 算力的指数级增长驱动了机器学习（ML）在过去 10 年中的发展[2]。
- 在微电子产业中领先数 10 年之后，美国将很快有约 90%的大容量、尖端集成电路生产被外包给东亚国家[3]。这意味着对于美国国防系统和更广泛产业都非常关键的尖端半导体生产几乎完全依赖他国，致使美国本土供应链变得脆弱，易受他国政府行为或自然灾害的影响。
- .当传统硅基芯片的结构出现边际性能提升衰退时，专用硬件、新颖封装技术（例如异构集成和三维堆叠）和新的设施类型将驱动未来的人工智能技术发展[4]。
- 随着军事和智慧社区（IC）继续把人工智能技术纳入到任务关键系统之中，对可信赖的微电子技术的需求只会保持增长[5]。

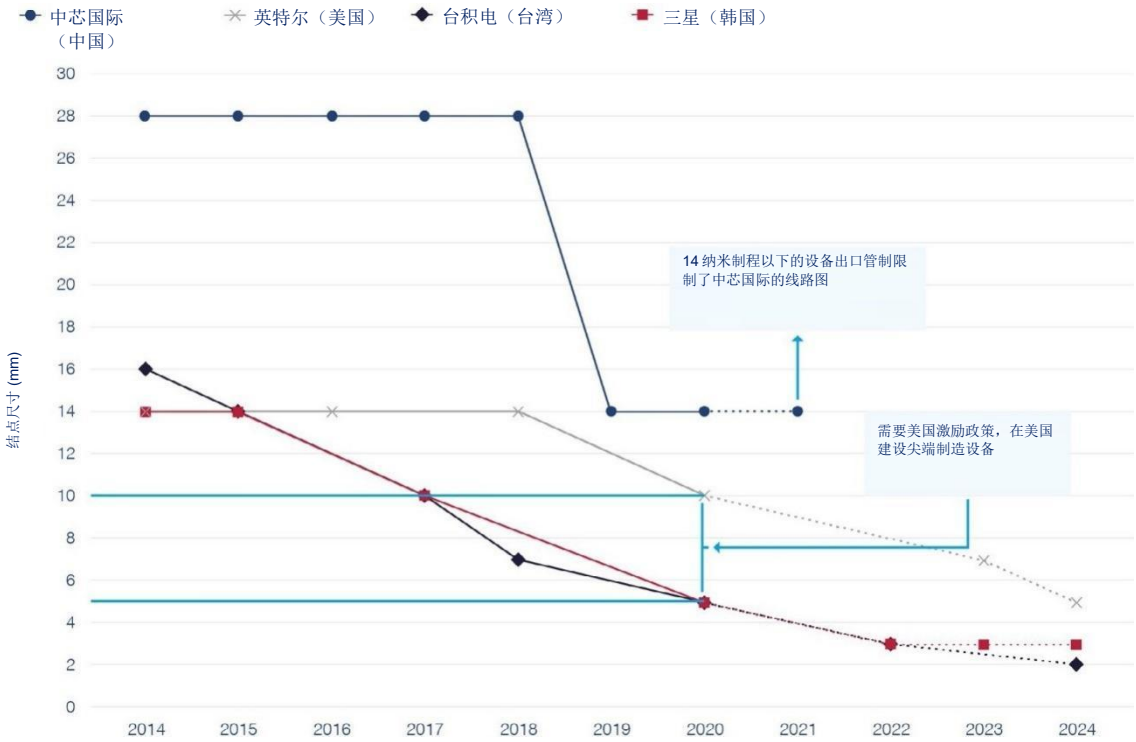
“对于美国国防系统和更广泛产业都非常关键的尖端半导体生产几乎完全依赖他国，致使美国本土供应链变得脆弱，易受他国政府行为或自然灾害的影响。”

美国作为微电子产业的先行者，一直被理所当然地视为半导体的领军人物。然而，美国却一直在失去其领先优势。虽然美国的大学和公司在半导体研发和芯片设计关键领域依然保持了全球领先，但半导体现如今已是高度全球化和竞争化的产业。台积电（TSMC）在半导体合同制造上全球领先，而韩国三星也在生产最顶尖的逻辑芯片[6]。台积电在基于 ARM 开发的芯片生产也处于领先地位，这种芯片正在成为移动设备、服务器和其他新兴技术关键应用的主流芯片结构[7]。为了在芯片生产上赶上竞争对手，实现芯片自主，中国正在史无前例地举国之力，期望在 2030 年打造一个世界领先的半导体产业。虽然就芯片生产而言，中国落后于总部在台湾、韩国和美国的企业，但中国芯片发展迅速[8]。虽然英特尔，这家美国的领先制造商，在芯片设计上保持了竞争力，但它在尖端芯片的生产上遭受挫折，可能被台湾和韩国的对手越甩越远，现今预测，2022 年英特尔在尖端产品制程上将落后两代，甚至更多[9]。这些以及其他令人担忧的趋势说明美国在微电子技术的领先优势正在被蚕食，尤其是在生产、装配、测试和封装上[10]。

敌对国家政府行为、自然灾害和其他事件容易导致电子产品的供应链中断，因此美国对半导体制造进口的依赖，尤其是台湾的进口，形成了美国经济和军事上的一个战略弱点。虽然美国在微电子研究、开发和创新方面有大量的专业知识，但还是受制于美国本土没有半导体制造设备，特别是没有顶尖半导体制造设备。如果当前趋势继续下去，美国将很快在芯片制造上掉队，并最终在半导体设计上被赶超。如果一个潜在对手长期在半导体行业压制美国，或突然完全切断美国获取尖端芯片的渠道，它就可以在战争中全方位占据上风。集美国政府、产业界和学术界之力共同发展本土微电子制造设施，将减少对进口的依赖，维持美国在技术创新的领先地位，支持就业，加强国防，改善贸易逆差，并巩固技术优势和战备优势，其中战备也是先进微电子产品的重要消费者。

“虽然美国在微电子研究、开发和创新方面有大量的专业知识，但还是受制于美国本土没有半导体制造设施...”

2014 年至 2024 年顶尖半导体的生产情况（按生产企业划分）



2014 年至 2024 年顶尖半导体的生产情况（按生产企业划分）

2021 至 2024 年的节点尺寸为预测值，反映了企业线路图。

节点尺寸体现的是第一年量产的制程。

因为 14 纳米制程以下的生产所需材料存在出口管制，此图没有显示中芯国际 2021 年以后的线路图。

美国要想在微电子再次领先，行政部门应该完成国家微电子领先战略规划，并落实执行。此外，国会应该为来自美国本土公司及其合作伙伴对实现本土制造而进行的投资予以40%的退税优惠，并在今后5年里为微电子研究、开发和基础设施另行拨付120亿美元。齐心协力让美国政府、私营部门和学术界迎难而上，重现美国在半导体上的优势。

目标：在最顶尖的微电子技术上，领先中国两代水平，并保持顶尖微电子在本土制造的多重来源

美国国家目标的当务之急应聚焦在微电子长期竞争所必需的环节和资源上：在最顶尖的微电子技术上，领先潜在对手两代水平，同时保持顶尖微电子在本土制造具有多个来源[11]。虽然美国的半导体设计和制造曾经在历史上领先中国至少两代水平，但这没有作为一个清晰的政策目标被提出。如今，虽然中国还没能超越美国，但其他国家及地区，如台湾和韩国，在顶尖半导体的生产上明显领先美国。这让美国在国防系统和更广泛产业都很关键的尖端半导体生产都依赖于他国。美国在半导体全球价值链上保持了局部优势地位，例如设计、电子设计自动化工具和半导体制造设备（SME）[12]。因此，美国应该直言不讳提出重获半导体领先地位的目标，集中政府、产业界和学术界，举全国之力，在不同领域争取领先，例如在美国已经落后的半导体生产领域，制定一定可以按照时间来追踪进度的标杆目标。要实现这一目标，委员会推荐举措集中在3个方面：

- 执行国家微电子战略；
- 通过激励在本土生产多重尖端设备来振兴本土微电子制造；以及
- 加强微电子研究。

除此之外，要推动美国在微电子技术上的领先，美国和盟友应利用高端半导体生产设备的定向输出管制（见本书第14章的描述），以保护现有技术优势，达到延缓中国半导体产业的进步。

要保持美国在 AI 技术上的领先，美国必须在微电子上处于领先

建议

实施*国家微电子战略*。美国缺乏国家微电子战略，以协调行政部门内部、产业界、学术界半导体政策、资金支持和激励政策。一个真正的国家战略应基于本委员会的工作，以及美国政府或代表美国政府之前所做的研究之上。该战略应综合国务院、国防部、能源部、商务部、财政部和其他相关机构各自的独立方法，推动国内研发和半导体制造专业知识，同时防止对竞争者的违法技术转移。最后，该战略应不断更新，形成协调一致的办法，以应对在微电子创新、竞争力和供应链整合中不断变化的挑战。

与委员会的建议一致，2021 财政年度的《国防授权法案（NDAA）》组建了国家科学和技术委员会（NSTC）的小组委员会，由高级政府官员组成，以制定微电子研究的国家战略，并监督其执行[13]。然而，国家战略的成功，关键在于该战略应被白宫列为重点，并要求 NSTC 小组委员会在 270 天内向总统提交国家微电子战略。

建议

重振本土微电子制造业。委员会的结论是，美国在微电子上过于依赖全球多元化的供应链，包括从潜在对手进口。而且，美国在产业基础上的差距，导致其可能失去获得可信赖的、有保障的、顶尖的、用于国家安全使用场景的半导体渠道的风险 。尽管忧虑重重，委员会对自去年以来美国在重振本土顶尖微电子制造发展而做出的各项努力，备受鼓舞。

> “...美国可能失去获得可信赖的、有保障的、顶尖的、用于国家安全使用场景的半导体的渠道。”

例如：台积电决定在美国开发先进设备，英特尔公开表示有兴趣和美国政府共同开发商用芯片晶圆代工厂。然而，这仅仅是起步，要实现最终美国本土拥有多个制造最尖端芯片企业的目标，美国政府依然任重道远。如果在美国本土没有几家拥有可以生产制造芯片的生产设施，美国产业界和国家安全都将面临竞争压力和供应链短缺的风险。这方面，近期最突出的进展包括《美国芯片制造法案》中一些与半导体有关的条款，已被纳入到 2021 财政年度《国防授权法案（NDAA）》中[14]。然而，这些项目的成功需要充足的财政拨款支持 ，而在 2021 财政年度并未获得拨付的支持资金，致使国会重点项目不清不楚。继续推动国会采取行动，设立退税机制，为高端微电子的本土化生产奠定基础，是美国保持领先中国两代水平的重要之举。具体而言，美国政府应该：

- **通过退税机制激励顶尖商用芯片制造本土化。**虽说这是《美国芯片制造法案》的内容，但国会还未通过立法，为半导体设施和设备提供 40%的退税[15]。现有美国的激励政策是降低资金成本、运营成本和税费，让晶圆代工的施工成本降低 10%到 15%。美国要想成为有竞争力的半导体生产市场，就需要一个与其他领先的半导体制造国同等数量级的税收优惠，例如韩国、台湾和新加坡，都实现了 25%到 30%的降本，差不多是美国现有水平的两倍[16]。激励政策是美国缺乏先进逻辑商用芯片晶圆代工厂的背后驱动力。弥补激励政策上的差距，并将激励政策放宽到包括来自盟国的企业，将激励美国企业在本土建立生产设施，并同时吸引例如台积电和三星这样的外国企业。此外，增加美国对高端 SME 的需求将为来自盟国（特别是日本和荷兰）的 SME 生产商创造新商机，这将使盟国政府更愿意将其出口管制政策保持与严格的美国政策一致，禁止出口此类设备到中国[17]。

> “...其他领先的半导体制造国及地区，如韩国、台湾和新加坡，都实现了 25%到 30%的降本，是美国现有水平的两倍。”

建议

*倍增联邦政府对微电子研发的资金支持。*使用硅基晶体管的传统结构生产芯片，芯片每提高一代，其边际性能提升都会衰减，因为它们已达到了物理法则作用的极限。在开发尖端硬件的设计阶段让美国领先潜在对手两代水平的比较优势，会随着硬件迭代差距的缩小而逐步消退。因此，在中期美国必须考虑多元化集成和其他新颖的硬件改进，以继续在创新上胜过竞争对手。长期而言，美国也必须继续以其投资组合方式在新材料和全新硬件方式上投资，例如量子和神经形态计算，在未来的微电子道路上大踏步前进。广泛投资和激励政策对于保持美国在半导体生产有关的其他优势领域继续领先十分重要，包括电子设计自动化工具和 SME。

美国政府有 4 大研究部门，主要通过产业界来聚焦中期和长期微电子技术突破：能源部、DARPA、国家科学基金会（NSF）和商务部。其现有的一批项目，例如 DARPA 的电子复兴计划，其目标研究领域方向正确，但应该再提升一个数量级，以实现必要的突破并保持美国的竞争力。额外资金支持不仅用于研究项目，且要投入到微电子开发必需的资金密集型基础设施中，包括国家半导体技术中心和 2021 财政年度 NDAA 中授权的先进封装原型活动。尤其是，封装技术的进步对于未来提升半导体的产能十分重要，因为各家公司的二维晶体管密度已达到了物理极限[18]。政府应该：

- **倍增联邦研究资金支持，引领新一代微电子技术。**委员会建议美国政府应大幅增加聚焦在微电子上的全范围研究工作。国会应为微电子研究再拨付 11 亿美元，为 2022 财政年度先进封装国家制造项目拨付 10 亿美元。在这些投资的基础上，资金支持水平应持续 5 年，总投资约 120 亿美元。这些资金金额与《美国芯片制造法案》和《2020 美国晶圆代工厂法案》所述的金额一致，但尚未拨付到位[19]。资金的去向与这些项目的现有聚焦领域及委员会之前的建议一致，应用于建设基础设施，并在前景看好的领域中谋求突破，例如超越极紫外光刻、三维芯片堆叠、光子、碳纳米管、氮化镓晶体管、领域特定硬件结构、电子设计自动化和低温计算的下一代工具。

“…封装技术的进步对于未来提升半导体的产能十分重要，因为各家公司的二维晶体管密度已达到了物理极限。”

第 13 章——尾注

[1] 参见 Dave Martinez 等人所著《人工智能：以往历史、现今发展和未来展望》，麻省理工学院， 27、n. 10（2019 年 1 月）https://www.ll.mit.edu/sites/default/files/publication/doc/2019 - 09/Artificial%20Intelligence%20Short%20History%2C%20Present%20Developments%2C%20and%20 Future%20Outlook%20 -%20Final%20Report%20 -%20Martinez.pdf（援引 Andrew Moore 等人所著《人工智能的开发和利用蓝图》，国际光学与光子学学会，《Persistent 公司 ISR IX（2018）的地对空多传感器互通性、整合和网络》）。

[2]近期机器学习（ML）的突破很大程度上是依赖于算力的提高。自 2012 年开始，在最大规模 AI 训练中使用的算力是以指数级进行增长的。参见 Girish Sastry 等人所著《*Older Headline Results 中使用的算力*》，OpenAI（2019 年 11 月 7 日） https://openai.com/blog/ai-and-compute/#addendum.

[3] 参见 Michaela Platzer 等人所著《半导体：美国产业、全球竞争和联邦政策》，国会研究处，12（2020 年 10 月 26 日），https://crsreports.congress.gov/product/pdf/R/ R46581.

[4] 参见 Sara Hooker 所著《硬件彩票》，arXiv（2020 年 9 月 21 日），https://arxiv.org/pdf/2009.06489.pdf.

[5] 参见 Gaurav Batra 等人所著《人工智能硬件：半导体公司的新机遇》，麦肯锡公司（2019 年 1 月 2 日），https://www.mckinsey.com/industries/semiconductors/ our-insights/artificial-intelligence-hardware-new-opportunities-for-semiconductor-companies.

[6]台积电（TSMC）已经开始生产 5 纳米的顶尖芯片，并计划在 2021 年年底生产 3 纳米芯片。三星也在生产 5 纳米芯片。英特尔预计至少到 2022 年才能生产 7 纳米芯片，并将生产外包给台积电。中国的企业在生产 12 纳米芯片。参见 Richard Waters 所著《英特尔期待新的关键技术谋求触底反弹》，金融时报（2021 年 1 月 14 日）https://www. ft.com/content/51f63b07-aeb8-4961-9ce9 -c1f7a4e326f0；参见 Mark Lapedus 所著《中国 加速先进芯片开发》，半导体工程（2020 年 6 月 22 日）, https://semiengineering.com/china- speeds-up-advanced-chip-development/；参见《5 纳米技术》，台积电（最后登录时间：2021 年 1 月 16 日），https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_ 5nm；参见 Debby Wu 所著《台积电 280 亿美元支出燃爆全球芯片竞争》,彭博新闻社（2021 年 1 月 14 日），https://www.bloomberg. com/news/articles/2021- 01-14/tsmc-profit-beats-expectations-as-chipmaker-widens-tech-lead；参见 Anton Shilov 所著《三星芯片晶圆代工新闻：正在生产 5 纳米芯片，四季度的高性能芯片出货量增加》，汤姆年代硬件（2020 年 11 月 1 日）, https://www.tomshardware.com/news/samsung-foundry-update-5nm-socs- in-production-hpc-shipments-to-expand-in-q4

[7] *ARM 公司和台积电宣布达成多年合作协议，在高性能计算机使用的 7 纳米鳍式场效晶体管加工技术上开展合作（2016 年 3 月 15 日）*，https://www.design-reuse.com/ news/39433/arm-tsmc-7nm-finfet.html.

[8] 参见 Michaela D. Platzer 等人所著，《半导体：美国产业、全球竞争和联邦政策》，国会研究处，2、25、27（2020 年 10 月 26 日）， https://crsreports.congress.gov/product/ pdf/R/R46581.

[9] 参见 Ian King 所著《英特尔的“惨败”预示着美国芯片时代的终结》，彭博新闻社（2020 年 7 月 24 日）, https://www.bloomberg.com/news/articles/2020 - 07-25/intel-stunning-failure-heralds-end-of-era-for-u- s - chip - sector.

[10]参见 Michaela D. Platzer 等人所著，《半导体：美国产业、全球竞争和联邦政策》，国会研究处，（2020 年 10 月 26 日），https://crsreports.congress.gov/product/pdf/R/R46581.

[11]委员会之前的报告提供了一系列初步建议，建议拓展可信赖的半导体获取渠道，增加微电子研发资金支持，对高端半导体生产设备出口到对手国家进行出口管制，并将尖端制造设备的生产本土化。

[12]参见 John VerWey 等人所著《美国半导体制造设备产业的健康和竞争力》，美国国际贸易委员会《产业办公室工作文件》（2019 年 7 月 1 日），http:// dx.doi.org/10.2139/ssrn.3413951.

[13] 参见 Pub. L. 116 -283, sec. 9906, William M. (Mac) Thornberry，《2021 年财政年度国防授权法案》，134 Stat. 3388（2021 年）。

[14]参见 Pub. L. 116 -283, sec. 9906, William M. (Mac) Thornberry，《2021 年财政年度国防授权法案》，134 Stat. 3388（2021 年）。这些条款授权了几个项目，都是委员会之前就认为对于美国在微电子技术上领先所势在必行的。尤其是，条款要求起草国家微电子领先战略，建立国家半导体技术中心，并为半导体初创公司和先进封装国家生产研究所成立孵化器，这些都与委员会之前的建议一致。

[15]这项激励政策将减少半导体企业在半导体生产设备和设施上的 40%税费，一直到 2024 年。而在 2025 年和 2026 年，税费优惠将分别为 30%和 20%

[16]参见 Antonio Varas 等人所著《政府激励政策和美国在半导体生产行业的竞争力》，BCG 和 SIA（2020 年 9 月），, https://web-assets.bcg.com/27/cf/9fa28eeb43649ef8674fe764726d/bcg-government- incentives - and - us- competitiveness - in - semiconductor- manufacturing - sep -2 0 2 0.pdf.

[17]关于 SME 出口管制的更多详情，请参见本文第 14 章。

[18] 参见《异构集成线路图：第一章：异构集成线路图综述和概要》，IEEE 电子封装协会（2019 年 10 月），https://eps.ieee.org/images/files/HIR_ 2019/HIR1_ch01_ over view.pdf.

[19]参见 S. 4130，《 2020 年美国晶圆代工法》，116 次国会（2020 年）。

第十四章　技术保护

提升美国技术保护能力

加强
管理能力

利用针对性
出口管制

加强投资筛
选信息披露

加强科研保护

保持美国创
新优势

技术创新胜过竞争对手的能力是美国所有技术领先战略的核心要素。推动研发、创业和人才发展则是成功的关键因素。然而，随着美国技术优势边际收益的收窄，同时其他国家不遗余力获取美国专门技术，美国也必须重新检视自己如何保护技术理念、硬件、公司和价值观。

美国面临着由国家引导的技术转移和被窃取的威胁，被针对的技术是人工智能（AI）和其他尖端、军民融合技术和基础研究。中国构成了最明显的挑战。中国发起了一项全方面的技术转移战役，旨在 2050 年成为一个“科技强国”[1]。这项战役有意针对美国的关键领域、公司和研究所[2]。中国通过绕过出口管制、与美国企业达成商业协议获取知识产权（IP）等方式，导致美国每年损失 3000 亿到 6000 亿美元[3]。这个数字只包含直接损失，不包括对美国经济长期的影响。同时，中国通过网络、人才招募项目和科研伙伴来利用开放的科研环境漏洞[4]。

“中国通过绕过出口管制、与美国企业达成商业协议获取知识产权（IP）等行为，导致美国每年损失 3000 亿到 6000 亿美元。”

俄罗斯也构成了明显的非法技术转移威胁，尤其是与国防应用相关的技术。俄罗斯在技术收集方面很激进，具有很强的能力。在往后十年里，俄罗斯都可能利用现有商业企业和学术机构以及传统间谍和网络间谍，继续构成技术转移的威胁[5]。

现代化出口管制和投资筛选

军民两用商业技术对国家安全越来越重要，因此为限制符合国家安全利益的商品或资金出现转移，美国应如何设计对应政策，这是今后十年内的决定性考验之一。出口管制不能只用于防止特定敏感设备转移给战略竞争对手，也可用于延缓竞争对手在可应用于国防的敏感技术领域的发展进度。

如果执行得当，延缓竞争对手发展的出口管控可以维持美国现有的长期国防优势。例如，美国对喷气发动机技术的出口管制已在长达 30 年的时间里成功阻止了中国政府想要在本土为军用飞机生产现代喷气式发动机的目标[6]。

“为限制符合国家安全利益的商品或资金出现转移，美国应如何设计对应政策，成为今后十年内的决定性考验之一…”

然而，按照现有的设计和使用情况看，美国出口管制和投资筛选程序对于人工智能竞争来说还不够完善。作为政策问题，投资筛选和出口管制是另一个年代中设计的，那时民用和军用技术的区分更清晰，美国和竞争对手之间很少有经济重叠。如今这两个条件都已不同。人工智能技术是军民两用的，而中美的新兴技术经济紧密相连，要实现可行、战略影响最大、经济成本最低的设计异常困难。虽然这些权衡和折中问题都是老生常谈，但也变得更为极端，因为人工智能技术的军民两用性质意味着对于国防最关键的一些单独组成部分在商用领域早已司空见惯。

与此同时，美国监管能力没有跟上技术发展的步伐，因为商务部、财政部和国务院都缺乏足够的技术和分析能力，对军民两用新兴技术进行有效设计和高效执行技术保护政策。国会近些年已经采取了一些重要措施，针对新兴技术产生的挑战采取技术保护制度，最引人注目的是《2018 年出口管制改革法案（ECRA）》和《2018 年外国投资风险评估现代化法案（FIRRMA）》[7]。然而，这两个法案通过 2 年后，两者的关键方面仍未落实完成，拖累了方案的实施，并让受影响的行业对此摸不清头脑[8]。

这些条件为政策制定者提供了一个艰难抉择，是选择缺少保护，以无法接受的水平让竞争对手获取敏感技术，还是选择过度保护，可能扼杀创新并伤害美国整体竞争力。有效管制的卡脖子目标必须为可以对竞争对手产生显著的滴入式战略成本，而对美国产业造成的经济损失最小。但此类卡脖子目标越来越难以琢磨。

建议

*明确声明引导未来美国军民两用技术保护政策的大原则。*美国必须采取更明智、更可预测的方式，对人工智能技术采取技术保护政策。政府应声明未来技术保护政策将存在四个大原则：

- 美国技术管控不能替代投资和创新。
- 促进和保护美国技术领先必须纳入到美国战略中。
- 在人工智能相关技术中行使出口管制时，美国必须头脑清醒，选择不同的卡脖子要点，并与盟国协调相关政策。
- 美国必须在人工智能相关技术上扩大投资筛选范围。

> **“针对人工智能相关技术行使出口管制时，美国必须头脑清醒，选择不同的卡脖子要点，并与盟国协调相关政策。”**

在技术层面，因为人工智能具有军民两用、广泛传播和大部分为开放源的性质，因此对管制制度构成特别挑战。而且，人工智能是建立在大量其他技术之上的新技术。鉴于人工智能技术无所不在本质，对人工智能算法的出口管制存在根本的风险，对管制范围界定不当会导致无意间限制了大量商用产品的出口，并对美国技术产业造成实质性伤害。有些人工智能算法很明显应该被管制，例如可以应用于战场的技术，此类软件大部分已经处于商业限制清单的管制下[9]。数据也是一类管制目标，特别是敏感的批量数据传输，但有效的数据管制面临和人工智能算法类似的困难[10]。而在人工智能堆叠技术上，人工智能堆叠硬件零部件则是传统出口管制最可行的目标。

建议

打造监管能力，充分执行 ECRA 和 FIRRMA。 美国也必须采取措施，加强其设计和执行有效技术保护政策的能力。近期内，商务部、财政部、国务院都应确保具有足够数量的技术专才，聚焦在技术保护政策上，更好利用由技术政策设计专家组成的外部顾问委员会。商务部也必须按照 ECRA 两年多前的授权，完善其必须管制的初步“新兴”和“基础”技术清单，并全面采用美国出口管制清单应对聚焦现代技术的国家安全挑战[11]。这么做是执行 ECRA 和 FIRRMA 的必要之举。最后，各部门和机构应该加快出口许可发放以及美国外国投资委员会（CFIUS）的申请程序，实现自动化审批，提高这些领域的效率并降低经济成本[12]。

建议

*要求来自美国竞争对手的投资者在向 CFIUS 提供交易信息披露时，范围应涵盖更广泛的一组敏感技术。*美国应修正 CFIUS 的授权和程序，使之更好应对与敏感、军民两用技术有关的现代挑战。具体而言，美国应提高在关键技术行业中对来自竞争对手的投资进行监管的能力，防止窃取知识产权，并确保对敏感技术的管制。美国竞争对手如今在大量投资美国的人工智能企业。从 2010 年至 2017 年，来自中国的投资者在美国人工智能初创公司身上倾注了超过 13 亿美元，人工智能也是中国公司对美国进行风险投资的最高技术领域[13]。然而，美国政府没有看透这些交易的内幕。CFIUS 负责对涉及国家安全风险的外国投资进行筛选，但只在美国公司生产出口管制产品时（很少公司有这个业务）要求公司披露投资情况[14]。因此，许多投资于美国人工智能公司的来自美国竞争对手国家的公司，就无义务就其投资情况向 CFIUS 披露信息。虽然 CFIUS 有权撤销此类交易，但在自身完善之前，他们被蒙蔽了双眼，也就导致了显著的技术转移风险。

对于来自中国和俄罗斯在敏感技术领域的投资，CFIUS 应提高投资企业的信息披露要求。国会应授权要求来自特别关注国（包括中国和俄罗斯）在国家安全敏感技术的投资，包括按照 CFIUS 定义的人工智能和其他“敏感技术”的相关应用，必须进行信息披露，让 CFIUS 有机会在完成交易前进行审查。敏感技术清单应与 ECRA 所规定的“新兴”和“基础”技术不同，范围更广，包括美国国家安全有关的关键产业，这些产业面临来自敌对资本、与国家安全相关的人工智能技术应用、半导体、通讯设备、量子计算和生物技术，以及 2025 年中国制造中确定的其他领域。

针对这组更广的技术提出的强制申请要求，只有在某些选定的美国竞争对手的公司申请时会限定范围，避免过度监管，保持资金的自由流动，更加洞悉中国和俄罗斯在关键技术上的投资，防止国家资助的知识产权窃取行为，并出于国家安全的考虑维持美国在人工智能技术上的领先地位[15]。

“对于来自中国和俄罗斯在敏感技术领域的投资，CFIUS 应提高投资企业的信息披露要求。”

建议

*利用在半导体生产设备（SME）上的针对性出口管制。*美国应尽可能使用出口管制，防止竞争对手获得可以增加其战略或军事优势的人工智能能力。要限制竞争对手的人工智能能力，美国出口管制的首要目标是生产高端芯片必需的复杂 SME。SME 是一个关键的卡脖子要点，同时出于以下理由也是一个具有吸引力的出口管制目标：

- 先进人工智能技术越来越依赖于高端计算能力[16]；
- 中国的高端半导体供应依赖于国际性公司；及
- SME 的制造是由美国及其盟友所专有和垄断。

“要限制竞争对手的人工智能能力，美国出口管制的首要目标是生产高端芯片必需的复杂半导体生产设备（SME）。”

在美国竞争对手中，中国是唯一一个试图培育出本土尖端微电子产业，打造大规模生产先进芯片能力的国家。延缓中国在高端半导体制造能力上的提升，有助于推迟中国形成具有在国防领域进行先进人工智能技术应用的芯片制造能力的进度。同时，按照本文第 13 章概述的内容，努力促进美国半导体的领先地位，将进一步实现委员会提议的美国政策目标：在尖端半导体设计和制造领先中国两代水平。然而，在通用半导体上的管制不大可能有效，鉴于大量国家都有能力生产此类芯片，如果单边实施管制，只会伤害美国半导体产业。

建议

*在 SME 的出口管制问题上，美国、荷兰和日本应协同一致。*生产 16 纳米制程及以下的芯片所需的复杂光刻工具，特别是极紫外（EUV）和氟化氩（ArF）浸没式光刻工具是最复杂和昂贵的 SME 类型。这些工具甚至比大型 SME 更专业，而美国、荷兰和日本控制了整个市场[17]。国务院和商务部应与荷兰及日本政府联手，在涉及高端 SME 尤其是 EVU 和 ArF 浸没式光刻设备方面，保持出口许可发放过程的协同一致，对出口到中国的此类设备许可采用默认推定为驳回申请的政策。这将延缓中国达成本地大规模生产 7 纳米或 5 纳米芯片目标的过程，并通过限制中国公司修复或更换现有设备的能力，以限制中国在 16 纳米及以下规程的芯片生产能力，该规程是委员会评估认为对先进人工智能应用用处最大的芯片[18]。

*利用针对性最终用途出口管制和报告要求，以防止将美国高端人工智能芯片用于侵犯人权。*美国必须采取措施，防止并阻止美国公司有意或无意中让人工智能技术被用于侵犯人权的目的。鉴于大部分人工智能设备的商用性质，其应用绝大部分是合法的，因此按照清单进行管制不适合。然而，最终用途和终端用户出口管制可能更有效。虽然最终用途管制不大可能防止战略技术转移至下定决心要得到这个战略技术的美国的竞争对手，但这种管制可以防止或阻止美国公司让某些设备的关键组件，尤其是高端芯片，被用于恶意的人工智能应用。

“美国必须采取措施防止并阻止美国公司有意或无意中让人工智能技术被用于侵犯人权的目的。”

商务部应禁止特定某些用于大规模监控应用的高性能人工智能芯片出口，强制出口此类芯片的美国企业证明买家将不会利用芯片侵犯人权，并要求该企业提供极度报告给商务部，列出所有销售到中国的此类芯片。这不是许可发放的要求，可能带来不确定性并导致延误，而是更像一个自我认证和产业的半监管报告。这个行动可以证明美国对人工智能使用的道德和责任感，促使美国公司之间采取道德行为，并让害群之马更难以利用先进的美国芯片用于邪恶目标[19]。

加强科研保护

美国科研企业应作为一个国家资产而被保护。对于科技领域存在的威胁，我们需要技术更娴熟的情报收集和分析能力，并将该信息广为传播。政府机构、执法机构和研究所可以随时获取所需工具和资料，围绕特定威胁和策略进行更细致的风险评估，形成透明度。政府和研究所都有责任保护核心价值观，抗击恶意行为。与志同道合的盟友和合作伙伴一同协调响应措施，围绕基础科研的公开性、科研诚信和知识产权保护加强规范。

加强科研过程的诚信将稳固开发科研的基础。然而，如果所用方法思虑不周，美国抗击应对技术转移的政策行为可能会伤害美国的竞争力和全球科技进步。美国保持和中国的尖端科研活动的合作有利于美国，美国也欢迎中国的博士级顶尖人才到美国大学学习并在毕业后以 85% 到 90%的比例留在美国。

建议

*构建保护美国科研环境诚信的能力。*国会应通过《学术研究保护法案（ARPA）》，并建立政府资助的科研安全卓越中心。根据 ARPA 立法成立一个专门负责科研保护的国家委员会，改进与外国威胁有关的开源情报的传播，并帮助政府和科研组织之间的学习和实践分享。

*在国际间，与盟友和合作伙伴协调科研保护措施。*科学和技术政策办公室、国务院和司法部应与盟友和合作伙伴协调，在与竞争对手有关部门进行的有害学术合作方面进行进一步的信息共享，并为减少这些行动带来的伤害而采取多边响应。此类外交手段应用来加强以开放基础科研的承诺为中心的全球规范，正如美国在《国家安全决策指针 189 号》文件中所正式确定的。为了这一原则,美国应尽力组建联盟承诺和研究完整性，这些价值观是创新和全球科研合作之根本，让不遵守这些价值观的国家靠边站。”

美国应尽力组建一个致力于本原则和科研诚信的联盟，这些价值观是创新和全球科研合作之根本，让不遵守这些价值观的国家靠边站。

------“美国应尽力组建一个致力于本原则和科研诚信的联盟，这些价值观是创新和全球科研合作之根本，让不遵守这些价值观的国家靠边站。”

*加强研究院的网络安全支持。*保护科研数据和知识产权不被网络窃取可能是最重要的措施，也是最容易实现的安全层面。当窃取训练数据或训练模型基本上就相当于拿到了最终产品，做好网络安全对人工智能技术上尤为重要。联邦资金提供机构应通过发放清晰指南、采取激励措施、分享最佳做法和资源，让研究所可以轻松实现基本水平的网络安全维护。

例如国土安全部（DHS）和联邦调查局（FBI）之类的机构，应加强支持信息分享的概念，并在网络威胁和侵入上提供及时可行的警示[20]。此外，政府应与商业云运营方进行中间协调，在大学科研团队和实验室开展让外国敌方高度感兴趣的科研时，保障数据存储的安全。

建议

*反对外国人才招募项目。竞争对手*的人工智能发展国家计划指示利用招募外国人才作为一个手段，形成人工智能专家的“高地”。这些存疑的项目在近年里越来越引人注目。这个项目不是通过具有吸引力的工作条件进行合法的科研人才竞争，相反，其很多方式是违反美国的研究诚信规范，违反信息披露规则，为技术转移创造了载体。

这个人才招募项目通常利用一种“兼职”招募方式，参与者可以保留在美国的职位，同时接受在竞争对手某研究所的岗位。签署合同时，通常因为要求专利贡献给竞争对手的研究所而导致冲突，即使该研究是在美国资助下进行的。

国会最近采取行动，在联邦资助科研的项目中要求强制性标准化信息披露，要求披露利益冲突、承诺冲突以及所有外部和国外的支持，以限制竞争对手人才招募项目的有害影响，对此我们表示赞赏。通过以机器可读的格式将资金申请和文档记录过程标准化和统一化，可以进一步强化国会行动的效果。

“信息披露和资金申请标准化的同时，每个科研资金支持机构都应采取强制和有充足资源支持的合规操作程序，形成执行信息披露政策和阻止害群之马的责任感。”

这些措施协同发挥作用，形成有效监管，自动检测欺诈行为，实现联邦科研资助机构的数据共享。信息披露和资金申请标准化的同时，每个科研资金支持机构都应采取强制和有充足资源支持的合规操作程序，形成执行信息披露政策和阻止害群之马的责任感。

建议

*加强签证审查，限制存疑的科研合作。*一些美国大学和研究人员无意中与竞争对手大学的研究人员开展科研合作项目，而这些大学与军方有紧密联系，从而让开展的科研项目直接帮助提升了军方和国家安全能力。我们发现访问学者或学生与军方的关系被故意淡化，或使用其原属学校的别名故意混淆隶属关系。

美国应该对来自指定关注国家的外国军事和情报组织下属研究所有关联的高等学位学生和研究人员的签证申请执行特殊审查过程，防止存疑机构的研究人员进入美国[21]。同时，要有充足资源确保收紧的审查过程，如果发现签证申请人有意隐瞒或未正确披露其军事或情报组织信息，则禁止发放签证。

第 14 章——尾注

[1]《国家创新驱动发展战略纲要》，中国共产党中央委员会和国务院（2016 年 5 月 19 日）（2019 年 12 月 11 日 CSET 翻译）https://cset.georgetown.edu/research/outline-of-the-national-innovation-driven-development-strategy/.

[2]*国家安全局副助理检察长 Adam S. Hickey 在第五届 CFIUS 和国家电信安全审查小组会议上的发言，美国司法部（2019 年 4 月 24 日）*， https://www.justice.gov/opa/speech/deputy-assistant-attorney-general-adam-s-hickey-national-security-division-delivers- 0.

[3]《中国的技术窃取行为是对美国最大的执法威胁》，FBI 评论，《卫报》（2020 年 2 月 6 日），https://www.theguardian.com/world/2020/feb/06/china-technology-theft-fbi-biggest-threat（引述于 William Evanina，国家反间谍与安全中心局长）。

[4] JASON 最近的研究发现“中国政府及其研究机构的行为与美国的科学道德观不一致，引发人们对外国势力在美国学术界的影响的顾虑……科研透明度方面存在问题，合作与联合中缺乏互惠互利，有承诺和利益冲突现象。”参见 JASON《基础研究案情》，MITRE 公司，39（2019 年 12 月）https://www.nsf.gov/news/special_reports/jasonse13curity/JSR- 19 - 2IFundamentalResearchSecurity_12062019FINAL.pdf.

[5] 2018 年，国家反间谍与安全中心宣称：“来自俄罗斯的对美国技术的威胁将在今后多年继续存在，莫斯科试图支持一个在贪污成风、国家控制、人才流失到海外求职的泥坑中挣扎的经济。”参见《外国网络经济间谍》，国家反间谍与安全中心，8（2018 年）， https://s3-us-west-2.amazonaws.com/cyberscoop-media/wp-content/uploads/201 8/07/26114025/2018ForeignEconomic-Espionage-Pub_ FINAL.pdf.

[6]参见 Robert Farley 和 J. Tyler Lovell《中国空军被其糟糕的喷气式引擎拖了后腿》，国家利益（2020 年 4 月 3 日），https://nationalinterest.org/blog/buzz/chinas-air-force-being-held- back-its-terrible-jet-engines-140252.

[7] 参见 Pub. L. 115 -232, Title XVII, Subtitle B, 132 Stat. 1636, 2208，修订版 Pub. L. 116 - 6, Division H, Title II, Section 205 《综合拨款法案》, 2019, 133 Stat. 13, 476; Pub. L. 115 - 232, Title XVII, Subtitle A, 132 Stat. 1636, 2174.

[8]特别指出，ECRA 要求商务部确认“美国国家安全必不可少、却未能控制的新兴和基础技术”，但时至今日商务部还没有确认本条款下任何一项技术。这让美国为保护其在高科技领域（包括 AI）的优势所采取的行动中留下了缺口，并对行业造成了不确定性。参见 Pub. L. 115 -232, Title XVII, Subtitle B, 132 Stat. 1636, 2208, 修订版 Pub. L. 116 - 6, Division H, Title II, Section 205《综合拨款法案》, 2019, 133 Stat. 13, 476.

[9] 参见 Carrick Flynn《人工智能出口管制的建议》，安全和新兴技术中心（2020 年 2 月 6 日），https://cset.georgetown.edu/research/recommendations- on-export-controls-for-artificial-intelligence/.

[10]关于与盟友和合作伙伴合作创建安全转移关键数据集标准，限制数据流向特定可信赖的国家方面，还有待进一步完善。关于本主题的更多详情，请参见本文第 15 章。

[11] 参见 Chris Darby 等人《缓解新冠疫情对经济的影响并维持美国在人工智能上的战略竞争力》，NSCAI，16（2020 年 5 月 19 日），https://www. nscai.gov/white-papers/covid-19 -white-papers/.

[12]第 14 章《行动蓝图》包含本建议的更多详情。

[13]中国在美国的风险投资在 2014 年后大量增加，但在 2018 年后有所停滞。然而，AI 领域依然是中国风投资金在

美国的最大投资领域之一。参见 Michael Brown 和 Pavneet Singh《中国在美国的风险投资，中国的技术转移战略》，国防创新试验单元（2018 年 1 月），https://admin.govexec.com/media/diux_chinatechnologytransferstudy_jan_2018_(1).pdf；参见 Adam Lysenko 等人所著《战略竞争新时代的中美风投资金》，（2020 年 1 月），https://publications-research.s3-us-west-2.amazonaws.com/RHG_Disruption_US+China+VC_January2020.pdf；参见 Mercedes Ruehl 等人《中国国家支持的基金在华盛顿当局设置障碍的情况下依然进入美国技术领域投资》，金融时报（2020 年 12 月 2 日），https://www.ft.com/content/745abeca-561d-484d-acd9 -ad1caedf9e9e.

[14]其结果是：CFIUS 的披露要求不成比例地影响了来自美国盟友的投资。2019 年，在 94 起 CFIUS 强制要求披露案例中，14 起来自日本，12 起来自加拿大，11 起来自英国，只有 3 起来自中国。递交国会的年报，CFIUS，33-36（2019 年） https:// home.treasury.gov/system/files/206/CFIUS-Public-Annual-Report- CY-2019.pdf.

[15] 这要求对 CFIUS 的授权立法进行修订，其文本草案见本文的立法草案附件。具体而言，它将修订《1950 年国防生产法》第 721（a）款内容（修订并编码为 50 USC § 4565[a]），并给予财政部新的授权，以变更强制申请要求，并厘清一份“敏感技术”清单，区别于出口管制清单。

[16] OpenAI 估计从 2012 年后，在最大型 AI 训练中使用的算力每 3.4 个月就翻一倍。参见 Dario Amodei 和 Danny Hernandez《AI 和算力》，OpenAI（2018 年 5 月 16 日） https://openai.com/blog/ai-and-compute/.

[17]荷兰企业 ASML 在 EUV 光刻工具上具有垄断地位，这是最先进的光刻技术，并且 ArF 浸没式光刻工具只有 ASML 和日本企业尼康能生产。

[18]瓦圣纳协定将所列可以生产 45 纳米或以下芯片的光刻设备作为管制产品。然而因为瓦圣纳协定不具有约束力，参与国没有义务必须把它当作法律规限来遵守。参见《军民两用商品和技术清单及军火清单》，瓦圣纳协定秘书处，72（2018 年 12 月），https:// www.wassenaar.org/app/uploads/2019/consolidated/ WA-DOC-18-PUB-001-Public-Docs-Vol-II-2018- List-of-DU-Goods-and-Technologies-and-Munitions-List-Dec-18-1.pdf.

[19]若交易涉及有关监控功能的产品或服务的外国政府终端用户，参考最近国务院提供的最佳做法指南。参见美国国防部《涉及有关监控功能的产品或服务的外国政府终端用户的交易，执行“UN 指导原则”的指南》，美国国防部（2020 年 9 月 30 日）https://www.state.gov/key-topics-bureau-of- democracy-human-rights-and-labor/due-diligence-guidance/.

[20]例如科研和教育网络信息和共享分析中心（REN-ISAC），REN-ISAC（最近登录时间 2021 年 1 月 2 日） https://www.ren-isac.net/.

[21]委员会建议将此作为 10043 号总统公告的更新内容，对于从事中国政府军民融合战略研究的中国公民暂停 F 或 J 签证。参见 85 Fed. Reg. 34353，《对来自中华人民共和国的特定学生和研究人员暂停非移民入境》，总统行政办公室（2020 年 5 月 29 日）https://www.federalregister.gov/documents/2020/06/04/2020 -12217/suspension-of- entr y- as- nonimmigrants- of- cer tain - students- and - researchers -from -the - peoples - republic.

第十五章　有利的国际技术秩序

与盟友和合作伙伴协调一致

制定国际科学与技术战略

调整国务院和美国外交政策

建立新兴技术联盟

发起国际数字民主倡议

实施全面的美国国家计划

培育国际新兴技术科研中心

美国自己单干是无法与专制对手雄心勃勃的全球技术相对抗。志同道合的国家必须携手共同推动建立基于规则之上的国际秩序，保护自由开放社会，并促进经济创新。专制对全球技术秩序的挑战包括了五个不同却又相互关联的方面：

- 美国和西方科技公司的全球市场份额所受挑战不断提升，美国及其盟友和合作伙伴的繁荣发展和全球经济地位也受之影响；
- 在使用中国开发的技术和中国建造的基础设施的国家，中国的影响力在不断提高，战略杠杆不断增强；
- 在可以轻易获得用于加强高压统治的数字工具的国家，专制强化的前景展望；
- 在一个政府以破坏自由价值观的方式利用数字工具的国家，对民主倒退的前景展望；
- 因为一批有影响力的国家有能力改变全球技术规范和标准，导致民主的盟友国家凝聚力有被破坏的威胁[1]。

“美国必须推动与我们的盟友和合作伙伴紧密协作的全面战略，确保人工智能的创新和应用符合对自由开放社会至关重要的价值观。”

美国必须推动与我们的盟友和合作伙伴紧密协作的全面战略，确保人工智能（AI）的创新和应用符合对自由开放社会至关重要的价值观。并且，美国也应与其最紧密的盟友和合作伙伴合作，建立使用人工智能工具时的道德和责任原则，维护国际技术标准的威信，促进数字市场，利用专业知识发展隐私保护技术，并分享方法和资源，防止专制者对数字基础设施和民主价值观的攻击。

为实现这些目标，委员会提议白宫指令工作由国务院领头，其他机构的分工为：

建议

*制定并实施科学技术国际战略（ISTS）。*在政府范围内以及与我们最亲密的盟友和合作伙伴，帮助制定人工智能和新兴技术政策；利用外来援助、专业技术和指南以及开发性金融工具；培育合作研发。ISTS 应作为国家科技战略的国际部分（见本文第 9 章）。ISTS 应围绕以下四个大方面：

- **建立一个新兴技术联盟**，包括盟友和合作伙伴，按照民主规范和价值观，促进新兴技术的设计、开发和使用；协调政策和投资，反对专制政权恶意使用这些技术；并提供具体、有竞争力的替代方案来取代中国制造的数字基础设施。
- 作为新兴技术联盟的一部分，与盟友和合作伙伴**发起一项国际数字民主倡议**，协调国际援助工作，开发、促进和资助对符合开放、保护隐私、安全和可靠的民主价值观、道德规范的人工智能和相关技术的应用。
- **实施一项全面的美国国家规划**，围绕技术标准、外来援助、开发性金融和出口管制，**支持国际技术工作**。
- 通过将美国国家人工智能研究所和多边机构的合作正式化，例如人工智能全球伙伴关系（GPAI），**提高美国作为一个合作研发工作的国际新兴技术研究中心的地位**，与关键盟友和合作伙伴一同在美国创建一个多边人工智能研究所（MAIRI），加快国际合作和人才交流。

建议

*成立一个新兴技术联盟。*美国应带领道同志合的国家组成的新兴技术联盟（ETC），或是作为一个更大的民主峰会的部分成果，或是单独的工作成果。紧接着 ETC 之后，要围绕以下几个关键区域组织好政策的同步：

- **建立并实施标准和规范**，支持民主价值观和建立安全、可靠和可信赖的技术；
- **在人工智能和数字基础设施上推动促进联合研发**，促进共同利益，造福人类；
- 通过反审查、反恶意信息作战、反人口贩卖和监控技术的非自由使用，**促进民主、人权和法治**；
- 通过授权协议、公用数据备案程序、在隐私加强技术上的合作投资，**探索促进在盟友和合作伙伴之间进行数据分享的方法**，并解决法律和法规方面的壁垒；
- **促进和保护技术创新**，尤其是通过出口管制、投资筛选、供应链保障、新兴技术投资、贸易政策、科研网络保护以及知识产权协同来实现；
- 通过分析劳动力市场的挑战、协调技能和证书要求，增加人才交流、联合培训和劳动力发展举措，来**培养人工智能相关人才**；
- **发起国际数字民主倡议**[2]。

建议

发起国际数字民主倡议（IDDI）。作为 ETC 的一部分，美国应联合盟友和合作伙伴共同发起 IDDI，协调国际援助工作，发展、促进和资助符合开放、私密、安全和可靠性的人工智能和相关技术的应用，

IDDI 战略

IDDI 对于确保全世界国家采取安全、可信和开放的数字生态系统非常关键[3]，为生态群落使用人工智能和数字技术赋能，增强民主，推动可持续发展，促进共享价值，如隐私保护、人权和法治。IDDI 继而为美国与志同道合的盟友和合作伙伴提供了一个机会，共同反对人工智能技术的独裁使用，尤其是为之提供了替代方案，替代用于非自由用途、危及民主社会之间的凝聚力并威胁集体安全的数字基础设施项目[4]。随着国际数字和通讯基础设施投资需求不断增长[5]，中国将继续使用数字技术发展以输出独裁主义并扩大影响，美国及其盟友和合作伙伴必须联合力量，协同战略，将政府协助工作的影响最大化，并加快私营部门投资，应对共同的挑战。

“...美国及其盟友和合作伙伴必须联合力量，协同战略，将政府协助工作的影响最大化，并加快私营部门投资，应对共同的挑战。”

建议

*实施一项全面的美国国家计划，支持国际技术工作。*就使用和支持美国外交政策工具而言，ISTS 应制定一个政府范围内的综合计划，包括技术和外来援助、开放性金融和出口管制，以推动 ETC、IDDI 和其他独立项目。如下所示，该计划应该包括调整国际技术标准的方法；协调和拓展国务院、美国国际开发署、美国国际开发金融公司和其他联邦机构的项目；使用针对性出口管制保持美国和盟友的关键技术优势，并促进技术的透明度和责任。要实现有意义的成果，则需要大量专用拨款[6]。

支持国际技术工作的美国国家计划.

科学技术国际战略（ISTS）特别小组
由白宫召集，由国务院、财政部、商务部、能源部、美国国际开发融资公司、美国进出口银行、千年挑战公司、国家科学基金会、美国国际开发署、美国贸易发展署和其他关键机构领导
制定并监督美国在政府范围内有关国际技术标准和国际数字发展工作的战略的实施

选择美国利益攸关者和拟担任的角色

国务院
对外政策领导和外交

高级领导人和巡回大使负责 ETC 和 IDDI 的制定和实施

完成历史工作，协调国际安全、经济政策、科技、人权和外来援助

通过美国使领馆促进技术外交

美国国际开发署（USAID）
数字开发和人道主义援助

通过数字战略优先实施数字开发

为更多美国数字开发项目提供资源、工具和专门知识

就国际技术标准提供建议

商务部

国家标准与技术研究院（NIST）
就国际技术标准提供建议

就技术标准协调政府机构间小组
加强与产业界的伙伴关系与合作
与关键盟友和合作伙伴就安全、可靠和可信赖技术协同标准一致

美国国际发展金融公司（DFC）
外国直接投资

在技术和数字基础设施领域扩大投资
增加混合金融交易，扩大规模

美国进出口银行（EXIM）
出口激励和资金援助

利用中国与转型出口计划加强美国技术竞争力
就激励和输出民主新兴技术提供建议

产业安全局（BIS）
终端用户管制

制定和协调终端用户许可发放政策和出口关注，促进民主价值观

美国贸易发展署（USTDA）
出口激励和技术援助

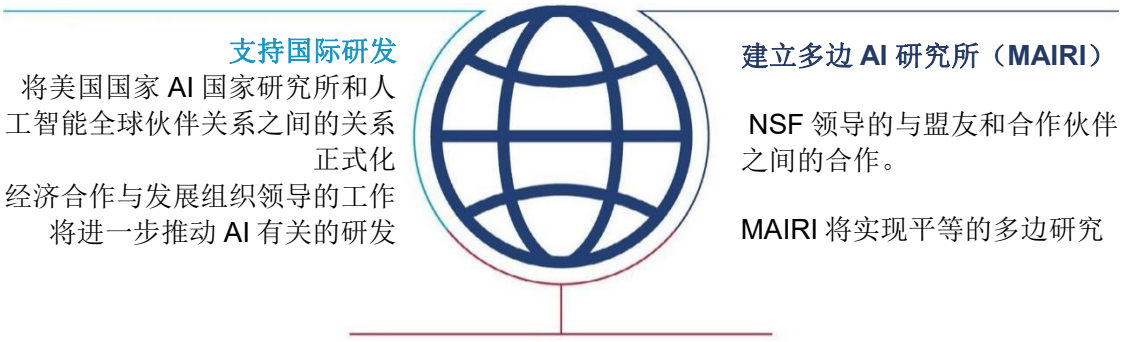
通过提高资金支持、培训、援助和试点项目支持美国新兴技术出口

建议

提高美国作为国际新兴技术研究中心的地位。美国必须通过进一步将自己确立为新兴技术及新兴技术演变的国际研究中心，促进与关键盟友及合作伙伴的人工智能合作与协同，从而维持自己在国际研发中的领先地位。这些工作将通过发展与民主价值观一致的数字技术和最佳做法，促进对 ETC 和 IDDI 的关键支持；提高美国对 GP 人工智能等现有和未来国际事务的贡献；并为美国和盟友特别是欧洲盟友提供了获取资源的光明大道，以弥补研发上的商用差距，并战胜关于跨境数据分享合作中出现的挑战。让美国成为国际新兴技术研究中心有三个要点：

“美国必须通过进一步将自己确立为新兴技术及新兴技术演变的国际研究中心，促进与关键盟友及合作伙伴的人工智能合作与协同，从而维持自己在国际研发中的领先地位。”

国际数字研究中心的组成



- **首先，美国应为关键的国际工作提供正式的科研资金支撑，例如，GPAI 和经济合作与开发组织[7]；**特别是通过国家科学基金会（NSF）的国家人工智能研究院[8]。国家人工智能研究院（由 NSF 和其他美国机构管理）和其他美国政府机构所承担的重要科研内容是一种非常重要的资源，应支持这些国际工作并推动美国与志同道合的合作伙伴的人工智能和数字技术目标。

- **其次，美国应该与关键盟友和合作伙伴共同建立多边人工智能研究所（MAIRI）。**MAIRI 将携手开发可以推动负责任、以人为本和保护隐私的人工智能/机器学习（ML）的技术，构建更好团体，让盟友吸纳人才和资源。MAIRI 将提供一个平等、多边的研究，加快基于志同道合国家优势的人工智能研发，为未来时代培养全球人工智能人员。MAIRI 将成为美国引领推动自由开放社会价值观、赢得全球技术竞争、促进人工智能创新和经济繁荣、开发造福人类的人工智能应用之关键。MAIRI 成员将一致拥护科研诚信原则、利用可信的基础设施和研究资源，并力争成为全球研究机构网络的一部分。NSF 应成为固定合作伙伴，但 MAIRI 的结构应允许其他联邦机构参与进来，例如国务院和能源部[9]。美国应资助初创公司，包括购买 MAIRI 位于美国的中心。

- **第三，美国应利用现有 O 类和 J 类签证项目，促进国外研究人员参与到联合项目中。**美国与盟友和合作伙伴之间持续的牢固合作，对于赢得这场科技竞争并促进志同道合国家的创新和创业至关重要。建立良好关系、交换观念和专门知识、激发未来合作，没有什么方式比肩并肩的共同研究更好[10]。

建议

*重新定位美国外交政策和国务院的角色，在数据时代赢得强国竞争力。*新时代外向型数字外交政策举措只是确保全球技术政策长期成功的方程式一部分。美国也必须对国务院进行聚焦内部的改革。国务院内部没有清晰的新兴技术政策或外交导向，这阻碍了国务院做出战略性技术政策决定的能力。同时也让盟友和合作伙伴感到困惑，在人工智能、5G、量子计算、生物技术或其他新兴技术有关问题上，他们常常搞不清哪个高级官员才是他们主要的联系人。

在人工智能和新兴技术竞技场形成有竞争力的外交策略是在强国竞争中的一项战略要务，国务院重新强化自身定位势在必行。美国必须重新设计国务院的内部架构、工作重点和文化，让美国外交适应数字时代，在技术、安全、商务和人权的交汇处，为美国谋求利益。支持这些工作，在美国外交上取得胜利，需要国会有针对性的拨款。

要为美国外交重新定位，委员会建议立即采取以下行动：

- **首先，国务院负责管理和资源的副国务卿（D/MR）**应负责围绕技术外交优先进行国务院的重新定位和重组。以往行政管理让 D/MR 负责战略重点和确保实施。D/MR 应围绕技术外交确定短期和长期规划方向，包括政策制定、协调和资源调配。D/MR 也应该在 ISTS 的监管和执行中负有领导职责。

“美国必须重新设计国务院的内部架构、工作焦点和文化，让美国外交适应数字时代...”

- **其次，国务院应该加速以人员和资源为重点，扩建新成立的网络空间安全和新兴技术局（CSET）。**CSET 在 2021 年 1 月初批准成立，由具备特使和协调官头衔的官员领导。CSET 是美国围绕与新兴技术相关的安全挑战的外交工作焦点，在国务院内部负责人工智能宣传[11]。国务院在国会支持下，必须确保 CSET 人手和资源充足。国务院要具备技术外交能力、改进各部门技术政策协调、经常把技术问题提高到高级领导人关注的层面，快速建立 CSET 举足轻重。要实现这些目标，国务院应对 CSET 所处位置进行评估，但也必须确保 CSET 的建立应该尽快完成。

- **第三，国务院应提高其在外国和美国技术中心的存在感，**在美国领事馆配备一些专业技术官员干部，以加强外交宣传，提高技术侦查、信息政策和外来援助工作。

- **第四，人工智能相关技术模块应纳入外交学院培训课程，**分不同级别，确保美国外交人员在新兴技术转型的环境中也具备领导才能。

- **第五，国会必须为国务院拨款用于紧急所需，**扩大美国外交团，支持聚焦人工智能和新兴技术的国务院关键项目，为美国谋求利益。

这些步骤都必不可少，但要在技术外交上为美国谋求更多利益还不够。最终，D/MR 的角色应转变为一个负责科学、研究和技术的永久性副国务卿（State/Q）。负责科学，研究和技术副国务卿将领导国务院的重组，将各办公室和各局合并，形成一个更健康、更协调的科学和技术外交策略和在强国竞争背景下的外来援助政策[12]。

第 15 章——尾注

[1]对盟友凝聚力的威胁也扩散到军方，不同或不兼容的数字系统构成了互通性挑战，或导致了美国军队在盟国行动的风险。参见 Daniel Kliman《为何美国需要一个数字发展基金会》，新美安全中心，2（2019 年 10 月 10 日）https://www.cnas.org/publications/ commentary/why-the-united-states-needs-a-digital-development-fund（长期来看，中国的数字投资将导致一些发展中国家变为美国军队的禁区，限制了美国军方进入的地理范围）。

[2]关于这些关键区域的详情，请参见第 15 章行动蓝图和相关附录。

[3]美国国际开发署的数字战略将“数字生态系统”定义为“利益攸关人、系统和赋能环境，三者共同为人和生态群落赋能，使用数字技术获得服务，彼此相交，或追寻经济机会。”这包括“一个理想的赋能环境和政策承诺、牢固和有弹性的数字基础设施、有能力的数字服务供应商和劳动力，以及最后赋能的数字服务终端用户。”《2020-2024 数字战略》，美国国际开发署，4（2020 年 6 月）https://www.usaid.gov/usaid- digital - strategy.

[4]中国政府的全球基础设施项目及其在私营部门广泛的政府影响力，使得中国企业为全球数百个城市提供监控和智慧城市技术，特别是在发展中国家，支持独裁政权，利于中国地缘政治胁迫和政府数据收集。参见 Hugh Harsono《中国的监控技术正监视着全世界的人口》，外交官杂志（2020 年 6 月 18 日）https://thediplomat.com/2020/06/chinas-surveillance-technology-is-keeping-tabs-on-populations-around-the-world/；Steven Feldstein 在中美经济和安全评论委员会上的证词，《关于中国在非洲的战略目标的聆讯》（2020 年 5 月 8 日），https://www. uscc.gov/sites/default/files/Feldstein_Testimony.pdf.

[5]为支持 G20 会议，全球基础设施中心已预测，在今后约 20 年间，全球通讯基础设施投资需要 8900 亿美元，按照当前趋势，供应将短缺 1000 亿美元。《预测基础设施投资需求和差距》，全球基础设施中心（最近登录时间：2021 年 1 月 13 日） https://outlook.gihub.org/.

[6]为美国机构和国会提供的详细建议，请参见第 15 章行动蓝图。

7 人工智能全球伙伴关系是 2020 年发起的，用以“培养有责任的 AI 开发植根于这些人权、包容、多元化、创新和经济增长中”。当前 GPAI 成员包括澳大利亚、巴西、加拿大、欧盟、法国、德国、印度、意大利、墨西哥、荷兰、新西兰、波兰、新加坡、斯洛文尼亚、韩国、西班牙、英国和美国，经济合作与发展组织和联合国教育、科学及文化组织则作为永久观察组织。尤其是，GPAI 通过以多方利益攸关者工作小组形式分享的研究和专业技术，弥补了“理论和实践的差距”。《关于 GPAI》，GPAI（最近登录时间 2020 年 1 月 6 日）https://www.gpai.ai/about/；《联合国教育、科学及文化组织作为观察者加入人工智能全球伙伴关系》，联合国教育、科学及文化组织（2020 年 12 月 10 日），https://en.unesco.org/news/unesco-joins-global-partnership-artificial-intelligence-observer.

[8] 《国家科学基金会的人工智能》，国家科学基金会（2020 年 8 月 26 日），https://www.nsf.gov/cise/ai.jsp.

[9]例如，虽然国务院可以提供外交政策专门知识，并支持与盟友和合作伙伴的数据分享和 AI 研究云的支持举措，但能源部可以提供在行业里或通过其国家实验室从事应用研究的关键专门知识，尤其是在高性能和量子计算上。

[10]这些要点的详细建议，请参见第 15 章行动蓝图。

[11]国务卿蓬佩奥批准了新网络空间安全和新兴技术局，美国国务院（2021 年 1 月 7 日）https://2017-2021.state.gov/secretary-pompeo-approves-new- cyberspace-security-and-emerging-technologies-bureau//index.html.

[12]国务院的这些组成部分应包括 CSET、海洋、环境和科学局、科技顾问办公室、网络事务协调人和分析中心的关键职能。

第十六章　相关技术

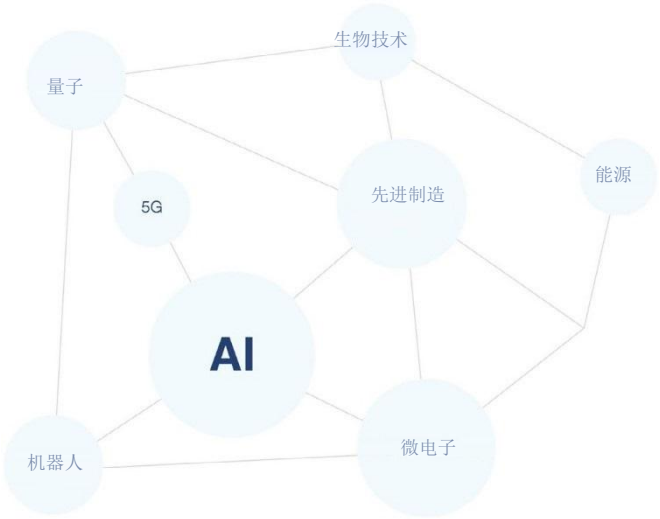
AI 及更多：新兴技术的确认和优先排序

委员会的工作最终又回到原点，结论是人工智能（AI）将在无形中改变我们生命的方方面面。然而，在人工智能技术上的领先并非是结局。对于保持美国技术领先的首要目标而言，人工智能技术领先是必要但不充分的条件。美国的战略面临一个问题：如何为人工智能和其他关键新兴技术进行优先排序，并支持基于跨技术优势并能放大优势的特定项目。美国必须从一系列更广的新兴技术竞争角度看待其在人工智能技术领先上所付出的努力，并支持美国的全面战略，从而维持关键技术上的领先地位。

“美国必须从一系列更广的新兴技术竞争角度看待其在人工智能技术领先上所付出的努力...”

在人工智能技术上的领先依赖于在一系列新兴技术上的领先，并进而推动在各新兴技术上领先。人工智能处于新兴技术群的中心，与其他新兴技术相互赋能[1]。例如，5G 和量子计算时刻准备着推动人工智能功能的新增长，而人工智能可以改变生物科学，形成引人注目的技术突破，将生物技术领域转变为整体经济竞争力的主要驱动力[2]。

人工智能处于新兴技术群的中心，与其他新兴技术相互赋能。美国政府必须将其在



美国政府:

为关键新兴技术定义并做优先排序，需要确认美国国家竞争力，包括美国在以下技术和平台的领先地位：

- 先进生物制造能力
- 量子芯片制造
- 5G 频谱共享
- 机器人软件
- 增材制造
- 储能技术

相关技术.

人工智能技术领先上付出的努力视为在每个新兴技术领域竞争的更广的战略的一部分。

中国正在实施一项全面技术领先战略。中国通过《2025 中国制造》在关键领域的战略投资威胁到美国的技术实力、经济繁荣和国家安全[3]。除在人工智能技术领域的投资之外，中国还试图成为量子、5G、生物技术和其他领域的世界领先者，并认识到在人工智能技术上的领先战略和其他技术上的领先战略之间的相辅相成关系。中国很清楚哪个技术应被视为国家优先重点，并在认为对整体技术领先所必需的广泛领域进行了大量投资。

而对于国家竞争力所涉及的核心技术，美国既没确认也没做优先排序。中国若具备了开发和利用如微电子、生物技术和量子计算的先发优势，美国将难以赶上。所以美国必须完成一份统一的清单，列出在 21 世纪构成国家竞争力之根本的技术，缺乏该清单导致美国各政府部门间的技术保护和技术推广的资金支持和政策方式各自为战。缺乏清晰的优先排序更难以有效管理在关键技术上的私营性质投资。在具有强大网络效应的关键领域，例如通讯行业，赢家通吃的趋势进一步加大了牌桌上的赌注，要快速发展领先技术平台，尤有甚者[4]。

确保美国在关键新兴技术平台的制造上处于领先定位，是国家竞争力的主要组成部分。要在新兴技术上维持国家竞争力，先决条件是对重点技术的科研内容确认并予以支持，但这还不够。美国也必须在这些技术的战略实体元素生产方面进行投资，建立扭转格局的平台，最大程度提高竞争力，并降低依赖性，防止因依赖性而构成国家安全弱点。此类投资通常非常昂贵，但战略方法不要求在国内生产所有的先进零部件，且战略投资将有可观的长期回报。几乎每个关键新兴技术领域都需要支持先进生产，包括半导体、量子计算、生物技术、通讯设备等，具体见以下建议。

实现技术领先需要在基本的数字基础设施上进行投资。要维持美国的整体技术领先，就无法忽视美国基本的数字基础设施状况。作为美国互连性基础的核心数字基础设施，也即所谓的高速互联网和通讯网络，其技术成熟度和范围远落后于许多发达国家[5]。此外，美国的实体基础设施大部分还是没有实现互联，美国没有一个城市进入全球连通性排前 10 强排名的智慧城市，前 30 强排名也只有一个[6]。让市民最大化地连接到数字经济中，确保他们具有必备的数字技能，和物理和数字世界更紧密连接，是推动未来增长的必由之路。提高美国实体资产的数字互联性不仅能提高其有效性和可靠性，也能产生新的数据源，实现人工智能技术在从电网管理和城市规划到交通等各个领域中的新颖而具有潜在革命性的用途。当美国思考其实体的基础设施的现代化将何去何从时，应优先完善其数字连通性，这将提供实质性的长远利益，并支持着美国的技术竞争力和国家安全。

确认和优先排序国家竞争力核心技术

虽然绝不会采取中国的中央计划和国家掌控的经济模式，但美国必须先制定更好的战略规划和预测，对新兴技术进行优先排序，以确保长远的竞争力。政府应该：

建议

*对要保障美国国家竞争力的关键新兴技术进行定义和优先排序。*白宫应公布一份美国领先且必不可少的关键和新兴技术清单，以之作为国家技术战略的一部分（参见本文第 9 章），并为每个领域制定详细实施计划，确定政府该如何与产业界携手推动技术领先，评估具体哪些是对于国家安全最为紧要的细分领域，并确定要营造所需投资环境所必需的法规措施或激励措施。

> “白宫应公布一份美国领先且必不可少的关键和新兴技术清单，以之作为国家技术战略的一部分（参见本文第 9 章）。”

这些规划应促进投资聚焦在可对美国技术领先产生乘数效应作用力的特定平台，确定关键卡脖子要点，在对美国产业影响最小化的同时封锁竞争对手，并提高供应链的弹性。列出清单、制定相关实施计划并落实行动，有助于政府、产业界和学术界对新兴技术经济竞争中最为重要的领域形成全国性共识。国会在考虑国家必须优先和扩大资源投入的方向时，全国性共识的结果应当具有重要借鉴意义，这也象征着对该产业的巨大需求。

美国政府内部也有许多类似清单，但却没人把各类清单集成为一个单一、权威的、具有战略眼光的文件，并提供详细的跟进行动计划，以确保美国的长久技术领先[7]。然而这些清单存在明显的内容重叠，这证明全国在关于国家竞争力最为紧要的技术方面，已涌现出共识。表 5 列举了委员会建议的初步技术清单，不论是否已经入选现有美国政府的关键技术清单，这些技术都是作为更广的技术领先战略的一部分。委员会建议白宫第一步先以行政命令方式认定这些技术为关键技术，然后指示政府部门和机构优先排序，并做相应协调安排。

美国政府关键技术清单						
NSCAI 提议的关键技术清单	2018 国防战略	国防部关键技术清单	商务部新兴技术的拟定规则的预先通知	总统科学技术顾问委员会的未来产业清单	S.383——无尽前沿法案	白宫关键与新兴技术国家战略
人工智能	✓	✓	✓	✓	✓	✓
生物技术	✓	✓	✓	✓	✓	✓
量子计算		✓	✓	✓	✓	✓
半导体和先进硬件	✓	✓	✓		✓	✓
自动化和机器人	✓	✓	✓		✓	✓
5G 和先进网络		✓		✓	✓	✓
先进制造			✓	✓	✓	✓
能源体系	✓	✓			✓	✓

采取行动推动对于美国技术领先和国家安全必不可少的技术和平台

在对于美国整体技术领先必不可少的新兴技术群达成共识之后，行政部门应对每个领域进行评估，并确定符合以下条件的特定平台：

- 具有战略意义和国家安全重要性的潜在应用；
- 可能对美国整体技术领先和竞争力有明显影响，无论是单独影响还是和现有的美国技术优势共同发挥影响；及
- 需要政府行动以刺激或保护其发展。

此类平台可能出于几个原因需要政府的支持。例如，在对国家安全具有战略重要性的领域，市场失灵可能会导致私营部门的投资不足。在其他情况下，只有联邦政府出于特定目的聚焦在私营部门、学术界和科研组织，才能抓住市场机会。政府必须通过提高资金支持、实施管理变革或采取其他措施根据实际情况区别对待，目的是为了根据实际环境促进技术革新和保护技术优势。

就上述几项战略技术中的关键技术平台，委员会已经提交了支持美国技术领先的建议。例如，本文第 11 章建议建设国家人工智能研究资源，以其为基础构建一个必不可少的平台，维持并延长美国在人工智能技术上的领先。此外，在本文的 13 章中，委员会提供了一系列关于促进美国在微电子技术领先的建议，包括激励在本地建成尖端商用芯片制造设备的具体行动。

以下建议是基于委员会之前的工作，并提供了美国政府可以采取的行动建议，以促进按照委员会评估认为最具有战略重要性的关键相关技术和平台领域中的技术领先，具体包括生物技术、量子计算、5G 和先进网络、自动化和机器人、先进制造和增材制造以及能源系统[8]。

生物技术

生物学已实现可程序化，而且人工智能识别编程优化方法的能力将导致生物技术出现变革性的突破。在新冠疫苗快速开发的过程中，人工智能具有决定性的作用，使研究人员可以在病毒完整基因序列被首次在线发布后，仅仅 2 天内就完成候选疫苗的基因序列[9]。用于医学影像的计算机视觉技术也实现了更准确和有效的诊断[10]。最近，一个人工智能网络在过去一年里为解决生物学最让人生畏的挑战中取得了实质性进展：从一个蛋白质的氨基酸序列中确定其 3D 形状[11]。此类工具结合合成生物学和基因编辑将变得更强大。

“生物学如今已实现可程序化，而且人工智能识别编程优化方法的能力将导致生物技术出现变革性的突破。”

对构成生命的基本成分进行深入研究，加快成果发现，并制造更先进的药物和材料，将共同增强人类的健康。随着人工智能技术不断推动生物科学的快速新发展，生物技术转变为整体世界经济的更大推动力，放弃在生物技术上的领先具有极大的战略影响，新冠疫情以清晰和明显的方式表明了这个事实。政府应该：

建议

优先发展先进的本土生物技术研发生态系统。美国应支持生物技术平台开发，将其作为国家生物经济学战略的一个部分，使研究人员利用人工智能技术推动新的生物学突破的能力最大化，并帮助从先进技术研究过渡到大规模的实际应用产品。这需要世界一流的生物数据资源，充分利用人工智能技术力量和生物制造平台快速实现分析突破带来的好处：

- **生物数据：**美国应该资助和优先打造一个世界一流的生物样本库，包括广泛的高质量的生物和基因数据组，让研究人员可以安全访问。“基因库”是领先的美国基因数据库，由国立卫生研究院管理，目前处于资金不足、利用率不足、管理不佳的境地。我们的目标是建立一个管理良好，便于研究人员访问和使用的基因数据库，包括大量和广泛的全体人类、动物和植物的基因组，对政府内部和私营部门开放专有数据集。基因库也应尽量包含相关表型的去识别化的元数据，包括对人类基因数据的强大隐私保护。这需要大量的资金支持。中国国家基因库，是中国对等的设施，由 BGI 集团（前身为华大基因），初期资金投入就要约 1.17 亿美元[12]。在美国建立这么一个实体，将提高生物技术的创新，并使之大众化，通过现有数据资源池化和促进新水平的人工智能赋能基因数据分析，同时减少美国研究人员为进行研究访问大型基因组数据库而对 BGI 或其他中国实体的依赖性。

- **生物制造：**美国应支持生物技术产业的多元化，并在现有的纵向整合模式外拓展该产业，鼓励开发多重的标准化的商用生物制造设施。要确保美国生物制造能力跟上人工智能对生物经济的变革性影响，这些工作不可或缺。让初创公司和实验室也有渠道使用高级生物制造工具，将让企业可以通过云服务快速设计新分子和材料，并立即下单制造。美国还应提供研发资金支持、采取激励措施，通过例如生物医学高级研究与发展管理局（BARDA）[13]之类的机构，支持先进生物技术制造措施加以适当管理，并扩大现有相关项目，例如 BioMADE[14]。国会应在未来健康有关的开支议案中优先资助此类项目。鉴于全球经济的实际投入中有 60%是通过合成生物学产生的，美国应在不断发展的生物制造技术上保持领先，这个需要即明确而又迫切[15]。

“...量子计算机在涉及机器学习和优化、物理系统模拟、敏感信息收集和传递有关的某类问题上具有胜过传统计算机的潜力。”

量子计算

根据摩尔定律推测，由于微芯片的物理极限，半导体生产商的创新步伐越来越困难，在下一代计算机硬件上处于领先地位对于美国在人工智能技术的战略技术上保持长期优势至关重要[16]。虽然在不久的将来，传统计算机还可能是从事日常计算任务最经济的方式，但量子计算机在涉及机器学习和优化、物理系统模拟、敏感信息收集和传递有关的某类问题上具有胜过传统计算机的潜力。例如，量子计算机可有效优化军事后勤，或发现武器系统的新材料[17]。在人工智能和量子计算的结合地带，每项应用都可以形成新式的国家安全威胁和机会。政府应该：

建议

*从基础研究转变为量子计算在国家安全上的应用，并激励本土制造。*美国在量子计算机研究上处于世界领先，但在国家安全的应用方面却可能失去领先地位。在认识到量子计算的进步可能推动人工智能技术进步的同时，美国必须为量子计算机建立可信赖的材料和零部件来源，投资于量子与传统的混合算法开发，并聚焦在国家安全的应用领域。通过国家人工智能研究资源提供传统和量子计算机的使用渠道，将促进利用了嘈杂的中间尺度量子计算机的量子与传统的混合算法开发。公开宣传量子计算机的特定政府使用场景，将对外表明国家安全应用市场的存在，并鼓励私营部门的进一步投资。

5G 和先进网络

5G 网络将成为人工智能平台之间的结缔组织，这意味着维护与可信赖和稳健的 5G 网络的通路也是在人工智能技术上保持整体技术领先的一个关键因素。华为在努力垄断全球 5G 网络，如今市场上在价格和质量上还没有单一一家供应商可以与之匹敌。鉴于在此情形下的紧迫性，美国应采取几个互补的方式同时在本土提升 5G 部署，并采用一个可靠的方案来替代华为。首先，所有各类措施都应包括对动态频谱共享的支持[18]。政府应该：

> **“扩大频谱共享对确保国防部进入作战效率所必需的频谱通道至关重要，同时要拓宽进入 5G 网络频谱的商用通道。”**

交易

*通过中频频谱共享支持和加速美国 5G 网络部署。*扩大频谱共享对确保国防部（DoD）进入作战效率所必需的频谱通道至关重要，同时要拓宽进入 5G 网络频谱的商用通道。扩展频谱共享和发放许可，以及批准开放额外中频端口以允许国防部和商用运营商同时使用，这需要多机构通力合作。通过这种组合方式，美国有很大几率可以加快 5G 的部署，这样部署速度才能保障支持人工智能的广泛应用。

自动化和机器人

自动化系统已经释放了全球市场的价值。在私营部门，产品已经实现了从专家咨询系统和无人驾驶到生产制造。在国防领域，自动化系统也有机会可以降低身处险境战士的数量，加快决策速度和质量，打造全新的军事实力[19]。先进机器人的软硬件设计和生产能力是自动化系统的关键所在。政府应该：

建议

*激励世界一流的机器人和自动化系统的软件平台开发。*随着各企业为不同使用场景和环境开发和定制机器人，自动化和机器人的未来可能体现为无限种类的形状和大小。美国追踪了一些国家的机器人和机器人硬件的部署情况，如中国、日本和韩国，发现自己必须提升某些领域的实力，例如机器人的材料设计和储能[20]。但是，美国在软件开发的专门知识，得以让自己为许多不同类的机器人硬件构建了一个世界级的数字平台。驱动机器人系统的软件将建立在几个根植于人工智能技术的核心实力上：机器人需要感知其环境、进行判断并在其所处环境中行动[21]。为这些核心实力提供尖端软件，是美国公司赢得软件平台市场，并推动下一次工业浪潮的时机[22]。为促进美国在自动化系统软件开发的领先，美国政府应该支持产业界的不断努力，补充提供基础研发、标准设置和国家标准与技术研究院（NIST）下智能系统部门所领导的数据共享项目[23]。政府也应采取激励措施以期尽早实现自动化的应用，在技术成熟领域为自动化系统创造市场，例如邮件分拣，这将为实现规模经济产生数据和经验，并应对临近市场[24]。自动化系统软件市场是和现有美国技术优势一致的具有战略重要性的领域。沿着上述发展路线，采取一个综合的、多管齐下的方式，会让美国产业界更有效地参与这个市场的竞争。

先进制造和增材制造

不论从维持成品的渠道角度出发，还是从作为技术创新的推动力角度出发，在本土生产高科技产品的能力都是国家安全的关键所在。关于渠道，美国必须力争在产业中实现自力更生，这是国家安全的关键，否则一旦出现旷日持久的冲突，要实现产业复兴就得长年累月了[25]。创新也会得益于技术设计和生产之间的紧密反馈循环，产品迭代会更加迅速[26]。在国防领域，这种关联尤其重要，将生产过程反馈到研发周期有助于技术从实验室走向实际军事行动。通过新技术例如增材制造对制造业的更长周期扰动也构成了对国家安全的威胁和机会。例如，增材制造可实现本地制造能力的突变，但也会因为可能让轻武器和其他军用产品的生产大众化，从而构成新的威胁[27]。政府应该：

> “不论从维持成品渠道的角度出发，还是从作为创新推动力的角度出发，在本土生产高科技产品的能力都是国家安全的关键所在。”

建议

*加快国防部各部门用增材制造生产遗存零件。*增材制造和 3D 打印具有改变制造业的潜力。可实现快速、高质量和复杂的生产，并且灵活度很高，可让 3D 打印机位于需求点附近，实现无库存即时生产[28]。虽然当前的增材制造技术还难以复制先进的传统生产技术，但人工智能技术已经展示了其大幅提高生产精度的能力[29]。联邦政府应主动支持可以促进增材制造技术开发的举措，并通过解决遗存零件的生产构成实际好处[30]。国防部应宣布在现役武器系统中确认所有可以通过增材制造和 3D 打印来生产遗存零件的目标，并在 2025 年实施。

能源系统

无论是为了保障军备、加快对国内危机的响应，还是为了维护经济良好运行，廉价和可靠的能源获取都是美国国家安全的关键所在。几乎每个领域都有能源投入，能源价格会直接影响经济产出，这也是美国国家竞争力的关键决定因素。而且，对外国的能源和技术依赖将导致美国处于一个容易被攻击的位置，尤其是当这些资源或技术被战略竞争对手所控制时。虽然美国在油气勘探、开采和加工方面位居世界前列，拥有足量的国内储备，但中国却在新能源方面遥遥领先，并在先进储能技术上大量投资，例如电池及其组分材料[31]。要保护竞争力，美国产业界应该在这些关键领域，大胆设定成本目标（按千瓦时和能量密度计算的价格）。在最具增长潜能的市场里，尤为如此，例如长续航固定储能设施和电动车的电池组[32]。政府应该：

*到 2030 年，开发满足美国市场需求的储能技术，并实现本地化生产。*开发新技术，以便更有效储存电能，随时随地可用，将推动输配电的技术进步，为美国提供经济和战略优势。为加快储能技术突破[33]，能源部制定了宏伟目标，要在 2030 年开发储能技术，并实现本地化生产，满足全美市场需求[34]。国会应充分资助联邦研发支出，并为实现 2030 年能源部储能大挑战线路图所需的商业化采取激励措施[35]。

“无论是为了保障军备、加快对国内危机的响应，还是为了维护经济良好运行，廉价和可靠的能源获取都是美国国家安全的关键所在。”

第 16 章——尾注

[1]认识到这个联系后，国会议题中纳入了 AI 和“相关技术”，因为它们在委员会的授权范围之内涉及到国家安全问题。

[2],委员会的第一个中期报告确定将生物技术、量子计算和 5G 作为 AI 相关的关键新兴技术。参见《中期报告》，50（2019 年 11 月） https://www. nscai.gov/previous-reports/.

[3]《2025 中国制造》包括以下领域：新一代信息技术、高端机加工和机器人、航天航空设备、海洋工程设和高科技船舶、先进轨道交通设备、新能源汽车、电动设备、农业机械、新材料和生物制药及高科技医疗设备。参见 Alice Tse 和 Julianna Wu 所著《为何“2025 中国制造”让特朗普总统怒不可遏》，南华早报（2018 年 9 月 11 日） https://multimedia.scmp.com/news/ china/article/made-in- China-2025/index.html.

[4]根据麦肯锡所述，“平台是驱动产品的后端技术实力，无论是由单独系统提供还是由多系统组合提供。”参见 Ross Frazier 等人所著《产品和平台：你的技术运营模式就绪了么？》，麦肯锡数字（2020 年 2 月 28 日） https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/products-and- plat forms - is -your-technology- operating- model - ready.

[5]按每百名居民的固定和移动宽带订阅数排名，美国在 37 家经济合作与发展组织（OECD）中排名第 18，而在平均宽带速度排名中排第 8。https://www.oecd.org/sti/ broadband/broadband-statistics/ （参见“渗透率和数据使用”表 1.2——每百名居民的固定和移动宽带订阅数（2019 年 12 月）和“速度”表 5.2 阿卡迈平均速度（2017 年 1 季度））。

[6] *智慧城市指数*，瑞士洛桑管理学院，（2019 年 10 月），https://www.imd.org/research-knowledge/reports/imd-smart- city-index-2019/#:~:text=The%20Top%2010%20smartest%20cities,and%20Dusseldorf%20(10th).

[7]例如，2020 年 10 月白宫发布了其国家关键和新兴技术战略，包括了各部门和机构确认的关键和新兴技术清单。该文件没有提供该技术对国家竞争力必不可少的理由，并缺乏对促进和保护美国在各技术上保持优势的具体方案。参见《国家关键和新兴技术战略》，白宫，A-1（2020 年 10 月）https://www.whitehouse.gov/wp-content/uploads/2020/10/National-Strategy-for- CET.pdf?utm_source=morning_brew. Michael Brown 和 Pavneet Singh 在其 2018 年的报告里，认为缺乏一个统一的关键技术清单，会伤害美国防止技术转移的能力，参见 Michael Brown 和 Pavneet Singh 所著《中国的技术转移战略》，美国国防创新实验单元，37（2018 年 1 月） https://admin.govexec.com/media/ diux_chinatechnologytransferstudy_jan_ 2018 _(1).pdf

[8] NSCAI 所提关键技术清单上 AI 技术以外的所有技术都包括于此，但半导体是在本报告的第 13 章单独说明，AI 赋能的生物技术要素也在本报告的第 1 章单独阐述。促进美国在生物技术、量子计算和 5G 的其他建议可参见行动蓝图第 16 章。

[9] 参见 Hannah Mayer 等《AI 技术让现代打败新冠近在咫尺》，哈弗商学院（2020 年 11 月 24 日），https://digital.hbs.edu/artificial-intelligence-machine-learning/ai- puts-moderna-within-striking-distance-of-beating-covid-19/; 参见 Noah Weiland 等《现代疫苗对新冠病毒具有高覆盖率》，食品及药物管理局发现，纽约时报（2020 年 12 月 18 日），https://www. nytimes.com/2020/12/15/health/covid-moderna-vaccine.html.

[10] 参见 Junfeng Gao 等《在医疗保健应用的计算机版本》，医疗保健工程杂志（2018 年 3 月 4 日），https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5857319/.

[11] 参见 Ewen Callaway《它将改变一切：DeepMind 的 AI 技术在解决蛋白质结构方面实现巨大飞跃》，自然杂志（2020 年 11 月 30 日）， https://www.nature.com/articles/d41586- 020 - 03348-4.

[12] 参见 Zhuang Pinghui《中国开发第一个国家基因库，旨在容纳数亿样本》，南华早报（2016 年 9 月 22 号），https://www.scmp.com/news/china/ article/2021623/chinas-noahs-ark-first-national-gene-bank-opens-shenzhen.

[13] 生物医学高级研究与发展管理局（BARDA），美国卫生与公众服务部（2019 年 12 月 4 日），https://www.hhs.gov/about/agencies/orgchart/aspr/barda/index. html.

[14] BioMADE 是一家和国防部公私合营的公司，由美国国家制造业创新网络运营，聚焦在建设一个可持续的生物工业生态系统，并提高美国生物工业竞争力。参见 BioMADE（生物工业制造和设计生态系统），美国国家制造业创新网络（最后一次登录时间 2021 年 1 月 9 日）https://www.manufacturingusa.com/institutes/ biomade.

[15] 参见 Michael Chui 等《生物革命》，麦肯锡全球研究所，43（2020 年 5 月 13 日），https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/the-bio-revolution-innovations-transforming-economies-societies-and-our-lives.

[16] 参见 Steve Blank《格罗方德半导体股份有限公司的败退真正意味着什么？》， IEEE 综览（2018 年 9 月 10

日），https://spectrum.ieee.org/nanoclast/semiconductors/devices/what-globalfoundries-retreat-really- means.

[17] 参见 Pontus Vikstål 等《利用量子近似最佳化算法解决排班问题》，应用物理评论，vol. 14, iss. 3（2020 年 9 月 3 日），https://doi.org/10.1103/ PhysRevApplied.14.034009; 参见 He Ma 等《在近期量子计算机上进行的量子材料模拟》，计算材料学（2020 年 7 月 2 日），https://doi.org/10.1038/s41524- 020 - 00353-z.

[18] *《5G 生态系统：国防部的风险和机遇》，国防部创新委员会（2019 年 4 月），*https://media.defense.gov/2019/Apr/03/2002109302/-1/-1/0/DIB_ 5G_study_04.03.19.pdf.

[19] *《自动化的夏日研究》，国防部科学委员会*（2016 年 6 月），https://dsb.cto.mil/reports/2010s/DSBSS15.pdf.

[20] 参见 Maximiliano Dvorkin 和 Asha Bharadwaj《哪个国家和产业用的机器人最多》，圣路易斯联邦储备银行（2019 年 11 月 7 日），https://www.stlouisfed.org/on-the-economy/2019/ november/robots-affecting-local-labor-markets.

[21]先进材料（例如生物成分）、人机界面和小而有效的电力供应是潜在创新的其他领域，可将机器人与 AI 和本章描述的其他相关技术相连。

[22]例如，核心实力也可能包括抓握物体的能力，这是机器人生产商 ABB 和几家欧美企业正在推动的技术。参见 Jonathan Vanian《工业机器人巨头联手一家成长的 AI 初创公司》，财富杂志（2020 年 2 月 25 日）https://fortune. com/2020/02/25/industrial-robotics-ai-covariant/.

[23] *智能系统部，*国家标准技术局（最后一次登录 时间 2021 年 1 月 6 日），https://www.nist.gov/el/intelligent-systems-division-73500.

[24]扩大自动化系统需求的一个特定举措是大幅扩大美国邮政服务的自主移动机器人试点项目，在 2025 年从 25 个分拣设施扩张到所有分拣设施。自主移动机器人和邮政服务， USPS 监察长办公室（2018 年 4 月 9 日）https://www.uspsoig.gov/sites/default/files/document-library-files/2019/RARC- WP-18- 006.pdf.

[25] *《关键技术可及性》，*国家学术出版社（2006 年），https://www.nap.edu/ read/11658/chapter/1；也参见《评估和加强美国制造和国防工业基地和供应链的弹性》，13806 号行政命令执行机构间特别小组（2018 年 9 月），https://media.defense.gov/2018/oct/05/2002048904/-1/-1/1/assessing-and-strengthening-the-manufacturing-and%20defense-industrial-base-and-supply- chain - resiliency.pdf（确认威胁美国制造和国防工业基地的 10 个风险原型）。

[26] *《美国先进制造技术领先战略》，*国家科学与技术委员会（2018 年 10 月），https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/10/Advanced- Manufacturing-Strategic-Plan-2018.pdf；参见 Gregory Tassey《重新振兴美国制造研发战略的理由和机制》，国家标准技术局（2010 年 1 月 29 日），https://www.nist.gov/system/files/ documents/2017/05/09/manufacturing_strategy_paper_0.pdf.

[27] *《竞争对手的 3D 机会》，*德勤（2017 年 8 月 22 日），https://www2.deloitte.com/us/en/insights/ focus/3d-opportunity/national-security-implications-of-additive-manufacturing.html.

[28] *《国防部用增材制造生产维修零件》，国防部监察长（2019 年 10 月 17 日），*https://media.defense.gov/2019/Oct/21/2002197659/-1/-1/1/DODIG-2020 - 003.pdf.

[29] 参见 Mark Anderson《有机器学习的 3D 打印更精确》，IEEE 综览（2020 年 2 月 19 日），https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/3d-print-jobs-news- accurate - machine - learning.

[30] 例如，2020 年 8 月，国防部为-52 喷气式发动机打印了第一个金属零件。参见 Kyle Mizokami《B-52 的老式发动机被 3D 打印升级》，大众机械（2020 年 8 月 10 日），https://www.popularmechanics.com/military/aviation/a33535790/air- force-3d-print-metal-part-turbofan-engine/.

[31] 参见 Robert Rapier《垄断世界矿物燃料生产的十大国家》，福布斯（2019 年 7 月 14 日），https://www.forbes.com/sites/rrapier/2019/07/14/ten-countries-that-dominate-fossil-fuel-production；国家排名，国际可再生能源机构（最后一次登录时间 2021 年 1 月 6 日）， https://www. irena.org/Statistics/ View-Data-by-Topic/Capacity-and-Generation/Country-Rankings.

[32] *《储能》，*美国能源部（最后一次登录时间 2021 年 1 月 6 日），https://www.energy.gov/ oe/energy-storage.

[33]储能领域包括广泛的技术基础，例如电池（传统和先进电池）、电化学电容器、飞轮、电力电子学、控制系统和用于储能优化和定型的软件工具。

[34] *《储能》*，美国能源部（最后一次登录时间 2021 年 1 月 6 日），https://www.energy.gov/ oe/energy-storage.

[35] *《储能大挑战：线路图》*，美国能源部（2020 年 12 月）， https://www.energy.gov/sites/prod/files/2020/12/f81/Energy%20Storage%20Grand%20Challenge%20Roadmap.pdf.