

# 为人工智能的未来做好准备

美国总统行政办公室

美国国家科技委员会

技术委员会

2016年10月

## 关于美国国家科技委员会

美国国家科技委员会是行政分支处理科技事务的首要机构，负责协调联邦政府各研发部门的科技政策。国家科技委员会的一项重要任务是为联邦的科技投资确定明确的国家目标，同时负责准备各种研发工具，以完成国家目标。国家科技委员会下辖 5 个委员会，包括环境、自然资源和可持续性委员会，国土和国家安全委员会，科学技术工程教育委员会，科学委员会和技术委员会。这些委员会负责监督侧重不同科技领域的分委员会和工作组。访问 [www.whitehouse.gov/ostp/nstc](http://www.whitehouse.gov/ostp/nstc)，以获取更多信息。

## 关于白宫科技政策办公室

白宫科技政策办公室依照 1976 年的《国家科学技术政策、组织和优先领域法》成立，职责包括就重大科技问题的相关政策及预算制定向总统提出建议；系统阐述总统的科技政策和计划；让联邦、州和地方政府与产业界和学术界的科研团体展开密切合作。白宫科技政策办公室主任同时还担任总统科技助理，并管理美国国家科技委员会。访问 [www.whitehouse.gov/ostp](http://www.whitehouse.gov/ostp)，以获取更多信息。

## 致谢

白宫科技政策办公室和总统行政办公室的工作人员以及其他部门和机构为这份报告的完成做出了贡献。由衷地感谢为之付出努力的每一位人士。

## 版权信息

这是美国政府一份涉及公共领域的工作报告，可以自由转载、复制和翻译；请注明报告出自白宫科技政策办公室，万分感谢。任何译本都需有免责声明，即翻译的准确性由译者，而非白宫科技政策办公室负责。任何译本的副本都需递交白宫科技政策办公室。该报告采用 CC0 1.0 通用授权，可在全球范围内使用。

美国总统行政办公室

美国国家科技委员会

华盛顿特区 20502

2016年10月12日

亲爱的同事们：

人工智能技术的进步开辟了新市场，同时也为医疗保健、教育、能源和环境等重要领域的进步带来了新的机遇。最近几年，机器在完成特定任务方面的表现已经超越人类，例如图像识别。专家们预测面向特定领域的人工智能技术将继续快速发展。虽然在未来20年内，出现能够在众多领域达到或超过人类水平的通用人工智能可能性极低，但机器将在越来越多的领域达到或超过人类的水平。

未来，人工智能将扮演越发重要的角色。为了帮助美国应对这种发展趋势，该报告对人工智能的发展现状、现有及潜在应用、人工智能技术进步引发的社会及公共政策相关问题进行了分析。就联邦机构和其他相关者如何采取进一步举措，报告给出了具体建议。另有一份文件与报告一同发布。该文件描绘了联邦政府资助人工智能研发的战略规划。未来几个月，政府还将公布一份后续报告，深入探讨采用人工智能的自动化技术对就业和经济产生的影响。

该报告由美国国家科技委员会的机器学习与人工智能分委员会完成，后者于2016年5月成立，负责跨部门协调工作，就人工智能相关问题提供技术和政策建议，监督各行业、研究机构及联邦政府的人工智能技术研发。美国国家科技委员会的科技委员会对报告进行了评估，认可报告内容。2016年，白宫科技政策办公室牵头组织了一系列公共宣传活动，包括与报告中提及的大学和其他协会联合举办的5场公共研讨会。该报告在这些活动之后完成。

2016年6月，白宫科技政策办公室发布《信息征询书》，共收到161条回复，9月6年公布了这些评论意见。与大数据一样，人工智能也将扮演“赋能者”的角色。该报告引用了政府此前公布的3份有关大数据的报告，并在它们的基础上完成。

只要产业界、民间团体和政府共同努力，推进人工智能技术积极面的发展，妥善管控风险和挑战，确保每一个人都有机会帮助建设人工智能增强型社会并享受所能带来的利益，人工智能便可在未来继续为经济增长贡献力量，同时也将成为一个改善世界面貌的宝贵工具。

谨致！

约翰·霍尔德伦

美国总统科技助理兼白宫科技政策办公室主任

梅根·史密斯

美国首席技术官

美国国家科技委员会

主席

约翰·霍尔德伦

美国总统科技助理兼白宫科技政策办公室主任

工作人员

阿弗亚·布鲁斯

白宫科技政策办公室执行主任

机器学习与人工智能分委员会

共同主席

埃德·费尔顿

美国副首席技术官

白宫科技政策办公室

行政秘书

塔拉·里昂

美国首席技术官政策顾问

白宫科技政策办公室

共同主席

迈克尔·加里斯

高级科学家

美国国家标准与技术研究院

美国商务部

以下联邦部门和机构均有人员在机器学习与人工智能分委员会任职。通过这个分委员会，他们共同努力，监督机器学习和人工智能（联邦政府内部、私营部门及国际社会）的发展，关注具有里程碑意义的重要人工智能技术的研发，通过联邦政府协调机器学习和人工智能相关知识与最佳实践的使用和共享，了解和询问联邦政府人工智能研发优先事项的进展情况。

商务部（共同主席）

国防部

教育部

能源部

卫生及公共服务部

国土安全部

司法部

劳工部

国务院

交通部

财政部

退伍军人事务部

美国国际开发署

中央情报局

总务管理局

国家科学基金会

国家安全局

国家航空航天局

国家情报总监办公室

社会保障管理局

总统行政办公室的以下机构也有人员在机器学习与人工智能分委员会任职

经济顾问委员会

国内政策委员会

行政管理和预算局

白宫科技政策办公室（共同主席）

副总统办公室

国家经济委员会

国家安全委员会

# 目录

|                      |    |
|----------------------|----|
| 执行概要 .....           | 9  |
| 导言 .....             | 12 |
| 人工智能简史 .....         | 12 |
| 什么是人工智能? .....       | 13 |
| 人工智能的现状 .....        | 14 |
| 对外宣传与报告撰写 .....      | 17 |
| 公共物品人工智能应用 .....     | 17 |
| 人工智能在联邦政府的应用 .....   | 19 |
| 人工智能与监管 .....        | 20 |
| 案例研究：自主车辆与飞行器 .....  | 21 |
| 研发与从业者培养 .....       | 25 |
| 对人工智能进展的监控 .....     | 25 |
| 联邦政府对人工智能研究的支持 ..... | 26 |
| 从业者的发展和多样性 .....     | 27 |
| 人工智能、自动化与经济 .....    | 30 |
| 公平、安全与治理 .....       | 30 |
| 安全和控制 .....          | 33 |
| 全球考量与安全议题 .....      | 34 |
| 国际合作 .....           | 35 |
| 人工智能和网络安全 .....      | 35 |
| 武器系统中的人工智能 .....     | 36 |
| 结论 .....             | 37 |
| 报告建议措施 .....         | 38 |
| 缩略语 .....            | 40 |

## 执行概要

未来，人工智能将扮演越发重要的角色。为了帮助美国应对这种发展趋势，我们对人工智能的发展现状、现有及潜在应用、人工智能技术进步引发的社会及公共政策相关问题，进行了分析。就联邦机构和其他相关者如何采取进一步举措，我们给出了具体建议。《国家人工智能研究与发展战略规划》将一同发布，该文件描绘了联邦政府资助人工智能研发的战略规划。

## 公共物品人工智能应用

人工智能和机器学习技术有助于应对世界上最严峻的挑战，扭转最严重的低效现象，进而改善人们的生活。这是人工智能和机器学习最让人感到乐观的地方。很多人将人工智能的发展前景与移动计算的进步相比较，后者对社会产生了颠覆性影响。对人工智能基础研究和应用研发的公共与私营部门投资已开始让公众收获巨大益处，例如医疗保健、交通运输、环境、刑事司法和经济融合等领域。借助于人工智能，各机构得以更快更有效地完成任务，进而提高政府的工作效率。

## 人工智能与监管

人工智能可以应用于很多产品，例如汽车和飞机。但这些应用必须受到政府监管，以保护公众免遭伤害，同时确保公平的经济竞争。将人工智能融入这些产品将如何影响相关监管措施？总体来说，在对人工智能产品采取监管措施以确保公共安全前，应对采用人工智能的风险进行全面评估，即在降低某些风险的同时会让哪些风险提高？如果一种风险处在当前监管部门的监管范围之外，就应该展开政策讨论，思考现行监管措施是否足以控制风险，或者是否需要采用人工智能。对人工智能产品进行监管将导致合规成本提高，或者阻碍造福公众的创新技术的发展和采用。鉴于此，决策者应考虑适当调整监管措施，在不影响安全或者市场公平的前提下，减少创新成本和障碍。

目前，人工智能产品监管挑战的例子主要集中于自动化车辆（例如自动驾驶汽车）和采用人工智能的无人飞行器。从长远来看，自动化车辆能够减少驾驶者失误，挽救很多生命，同时还能提高老年人和残疾人的行动能力；无人飞行器则能带来很大经济效益。但随着这些技术进行测试，并开始走向成熟，我们必须采取措施以确保公共安全。现在，美国交通部正利用原有经验调整监管措施，打造安全空间和供试验的测试基地，同时与产业界和民间团体合作，根据实际效果制定监管措施，以便在安全运作案例不断增多情况下，为更多应用开绿灯。

## 研发与从业者培养

在推动人工智能发展过程中，美国政府能够扮演一个重要角色——在加大研发投入的同时培养技术过硬的多元化人才队伍。一项联邦政府资助人工智能研发的战略规划与这份报告一同公布。该规划探讨了联邦政府在研发中的角色，指出了存在机遇的领域，同时就如何协调研发工作以实现效益最大化，和打造高素质的人才队伍给出一系列建议。

鉴于人工智能的战略重要性，联邦政府应密切关注全球人工智能技术的发展情况以做到早期预警——及时发现其他地区出现的重大变化，并适时调整美国的政策。人工智能的快速发展大幅提高了对拥有相关技能的人才需求，以推动该领域的前进步伐。在这个遍布人工智

能的世界，公众需要具备相关知识，才能读懂、使用和解释数据，才能参与政策讨论，探讨受到人工智能影响的领域。人工智能相关知识和教育培训越发成为联邦科学、技术、工程学和数学教育计划的强调重点。此外，人工智能教育也是美国总统倡导的“全民计算机科学行动计划”的一部分。该行动计划旨在让从幼儿园到中学的所有美国学生学习计算机科学，进而具备当前这个技术推动的世界所需要的计算思维能力。

### 人工智能的经济影响

人工智能短期内的经济影响主要体现在任务自动化上，即让以前无法自动化的任务实现自动化。虽然能够提高生产率和创造财富，但自动化也会对某些工作岗位造成不同程度的负面影响。借助于自动化技术，对某些技能的需求将减少，对人工智能相关技能的需求则会增加。白宫经济顾问委员会进行的分析显示，自动化对低薪工作岗位造成的负面影响最大，人工智能型自动化技术将拉大高教育程度和低教育程度工人之间的差距，加剧经济不平等。不过，适当的公共政策能够应对这些风险，确保工人获得相关培训，在辅助自动化的岗位上获得成功，而不是与自动化展开竞争。此外，合理的公共政策还能确保人工智能创造的经济效益得到广泛分享，让人工智能成为一个值得信赖的全球经济新时代的引领者。

### 公平、安全与治理

随着人工智能技术得到更广泛的应用，众多技术专家、政策分析师和伦理学家对普遍采用人工智能可能带来的意外后果表示担忧。如果利用人工智能做出与人员有关的决策——往往取代人为因素驱动下的繁复流程做出的决定——如何确保公平公正，又如何进行问责，这些都是忧虑的所在。白宫 2014 年发布的报告《大数据：抓住机遇，坚守价值》以及同年由总统科学技术咨询委员会提交给总统的报告《大数据与隐私：基于技术角度的分析》也提到了这些忧虑。对透明度的忧虑不仅聚焦数据以及采用的算法，同时也聚焦如何对基于人工智能的决策做出解释。人工智能专家警告称，了解和预测先进人工智能系统的行为存在固有挑战。

利用人工智能控制真实世界的设备将引发安全忧虑，尤其是在系统暴露于复杂人类环境的情况下。确保人工智能安全的一个重大挑战是研发出强大的系统，能够从实验室的“封闭世界”安全过渡到会发生无法预测事件的外部“开放世界”。对无法预知的情况做出响应是实现安全操作的难度所在。研制其他安全苛求系统和建造基础设施——例如飞机、发电站和车辆——获取的经验让人工智能实践者获益良多，帮助他们了解如何进行检验和确认，如何打造一个技术安全案例，如何管控风险以及应以何种方式让利益相关者了解风险。在技术层面，公平和安全方面的挑战相互关联。两种情况下，实践者都希望避免意外行为，同时获取所需证据，让利益相关者相信不可能发生意外故障。

对人工智能实践者和学生进行伦理教育是一项必不可少的举措。理想的情况是，每一个学习人工智能、计算机科学或者数据科学的学生都参与伦理和安全相关课程，并展开讨论。然而，仅仅强调道德还远远不够。道德教育虽能让从业者意识到他们对所有利益相关者负有责任，但与此同时，这种教育还应辅以技术工具和手段，通过技术方面的努力防止出现无法接受的后果，最终将良好的意图付诸实践。

## 全球考量与安全议题

在国际关系和安全的诸多领域,人工智能引发了一系列政策问题。在最近的国际会议上,人工智能成为了热门讨论话题,国家、多边组织和其他利益相关者开始评估人工智能所能带来的益处和挑战。他们之间的对话和合作有助于应对共同挑战,推进人工智能的研发,最终利用人工智能造福人类。

现在,人工智能在网络安全领域得到重要应用,未来将在防御性和进攻性网络策略方面扮演越发重要的角色。设计和操作安全系统需要耗费大量时间,同时还需要大量专家的参与。让这份专业工作部分或者全部实现自动化能够以极低的成本,大幅提高系统和应用的安全性,同时提高国家网络防御的敏捷性。利用人工智能同样有助于保持快速反应,迅速探测不断变化的威胁并做出响应。

人工智能在武器系统上的潜在应用提出了若干具有挑战性的问题几十年来,美国一直都在把自动化引入部分武器系统。这些技术改进有助于在使用这些武器系统时实现更精确、更安全、更人道的效果。不过,放弃人类直接控制武器系统的做法将不可避免地带来诸多风险,同时也会引发法律和伦理问题。

进一步结合自动和半自动武器系统与美国国防计划和部队结构的关键是继续确保所有武器系统(包括自动武器系统)的使用符合国际人道主义法。此外,美国政府应继续采取适当的措施控制扩散,与合作伙伴和同盟国共同制定关于武器系统开发和使用的标准。美国积极参与当前有关致命自主武器系统的国际讨论,同时也预见到这种武器系统仍将是未来国际讨论的一大焦点。目前,美国政府机构正就自主和半自主武器制定一项涉及整个政府,同时符合国际人道法的政策。

## 为未来做好准备

如果产业界、民间团体、政府和公众紧密合作,支持人工智能的研发,在重视挖掘人工智能潜力的同时妥善管控风险,这项技术便可成为经济增长和社会进步的强力推动力。

在此过程中,美国政府可以扮演多个角色。它可以召集会议,讨论重大问题,帮助确定公共讨论议程;可以在人工智能发展过程中监督安全和公平;制定合理监管举措,在鼓励创新的同时保护公众;提供公共政策工具,确保因人工智能改革的工作方式和手段提高效率,同时避免特定产业的工人遭受负面经济影响;支持能够造福公众的人工智能基础研发和应用;为培养技术过硬的多元化人才队伍提供支持。美国政府可以利用人工智能,以更低成本和更高效率提高公共安全。很多公共政策领域——从教育到经济安全,再从国防、环境保护到刑事司法——都会随着人工智能的进步迎来新机会,同时也会面临新挑战。美国政府必须不断提高自己的能力,了解和应对这些变化。

随着人工智能技术的继续进步,从业者必须确保采用人工智能的系统可控、开放、透明、可理解,且能够与人类有效合作。此外,它们的操作还需符合人类的价值观和目标。研究人员和从业者已然提高了对这些挑战的关注,未来也应继续予以重视。研究机器智能可以帮助

我们进一步了解人类智能。如果使用得当，人工智能可以“放大”我们的智能，帮助我们开创一个更美好、更智能的未来。

报告第 38 页将列出完整的建议列表。

## 导言

人工智能有助于应对人类社会当前面临的一系列严峻挑战。在世界范围内，智能车辆每年可挽救数十万生命，同时提高老年人和残疾人的行动能力。智能建筑可以节省能源和减少碳排放。精准医疗能够延长寿命和提高生活质量。智能政府能够更快速地为民众提供服务，提高服务的准确性，在危急时刻更好地保护公众，同时还能节省资金。人工智能加强型教育能够帮助教师让每一名学生获得良好教育，让他们过上安乐幸福的生活。在发展人工智能过程中，如果在关注其益处的同时谨慎管控风险和挑战，这项技术能够在很多方面造福人类。

美国一直处在人工智能基础研究的前沿。从联邦研究经费到政府实验室，人工智能发展史的大多数阶段都有美国政府支持的身影。联邦政府对非机密人工智能研发的支持主要通过，网络与信息技术研发委员会、国防高级研究计划局、国家科学基金会、国立卫生研究院、海军研究办公室和高级情报研究计划署。国家战略计算行动计划、大数据行动计划、使用先进革新型神经技术的人脑研究计划等国家级大型研究计划，也为人工智能的进步做出了间接贡献。人工智能技术当前以及未来所能带来的效益非常巨大，能够提高美国的经济活力和生产力，同时改善民众福利。与报告一同发布的一份文件描绘了联邦政府资助人工智能研究的战略规划。

未来，人工智能将扮演越发重要的角色。为了帮助美国应对这种发展趋势，我们对人工智能的发展现状，当前及潜在应用，人工智能技术进步引发的社会及公共政策相关问题，进行了分析。就联邦机构和其他实践者如何采取进一步举措，我们给出了具体建议。

## 人工智能简史

自从进入电子计算时代，让计算机拥有类似人类的智能一直是计算机专家的梦想。虽然“人工智能”这个词直到 1956 年才出现，但这一领域的源头至少可追溯到上世纪 40 年代。在 1950 年的著名论文《计算机器与智能》中，阿兰·图灵阐述了人工智能的想法。这篇论文提出了一个疑问——机器能够思考吗？此外，文中还提出了一项用于回答这个问题的测试，同时还指出机器可以像小孩子一样，通过编程从经验中不断学习。

在随后的几十年，人工智能领域几经沉浮，因为研究过程中遭遇的一些问题的难度远远超出预计，还有一些难关被证明无法借助当时的技术攻克。直到上世纪 90 年代晚期，人工智能的研究才开始进入加速阶段，此时的研究人员更关注人工智能的次要问题以及如何利用人工智能解释现实世界的问题，例如图像识别和医疗诊断。1997 年，IBM 的国际象棋电脑高手“深蓝”战胜世界冠军加里·卡斯帕罗夫，标志着人工智能技术取得了一项具有里程碑意义的成功。其他重要突破包括国防高级研究计划局的“学习和组织能力的认知助手”计划，

它促成了苹果语音助手 Siri 的问世；IBM 的超级计算机“沃森”在美国电视智力节目《危险边缘》中取得冠军；以及国防高级研究计划局在 2000 年代举办的自动驾驶汽车挑战赛获得惊人成功。

当前的人工智能发展浪潮以及对这项技术的狂热开始于 2010 年左右，主要归功于 3 个相辅相成的因素——电商、传统企业、社交媒体、科学界和政府提供了可用的大数据；大数据为大幅改进机器学习方式和算法提供了宝贵素材；以及大数据依赖的拥有强大运算能力的计算机。在这一时期，人工智能的发展速度令专家们惊讶不已。例如，在一场颇为流行的图像识别挑战赛中，最先进的人工智能技术让识别错误率从 2011 年的 26% 降至 2015 年的 3.5%。相比之下，人类的识别错误率大约在 5% 左右。

与此同时，产业界也不断提高人工智能的研发投入。2016 年，谷歌首席执行官桑达尔·皮查伊表示：“机器学习（人工智能的一个分支）是一项颠覆性的核心技术，促使我们重新思考我们做一切事情的方式。我们将这项技术应用于我们的所有产品，包括搜索、广告、YouTube 或者 Google Play。我们还处于发展初期，但你们终会看到我们将机器学习系统应用到所有这些领域。”毫无疑问，这种对待人工智能的态度将对软件的研发产生深远影响。科技业的很多首席执行官纷纷“拥抱”人工智能，其中就包括 IBM 的罗睿兰。罗睿兰表示 IBM 已将赌注压在人工智能上面。

### 什么是人工智能？

当前还没有一个被实践者普遍接受的人工智能单一定义。一些人将人工智能宽泛地定义为一种先进的计算机化系统，能够表现出普遍认为需要智能才能有的行为。其他人则将人工智能定义为一个不管在真实环境下遭遇何种情况，都能合理解决复杂问题或者采取合理行动以达成目标的系统。

对于人工智能的问题解决方案，专家们给出了不同分类。一部流行的人工智能教科书进行了以下分类：（1）能够像人类一样思考的系统（例如识别架构和神经网络）；（2）行为与人类相似的系统（例如借助自然语言处理通过图灵测试，知识表达、自动推理和学习）；（3）能够进行理性思考的系统（例如逻辑运算，推理和优化）；（4）行为理性的系统（例如通过感知、规划、推理、学习、交流、决策和行动来实现目标的智能软件主体和类人机器人）。风险投资家弗兰克·陈将人工智能的问题空间分解成五大项：逻辑推理、知识表达、规划与导航、自然语言处理和感知。人工智能研究员佩德罗·多明戈斯根据人工智能研究人员采取的方式，将他们分为五大流派——根据抽象符号进行逻辑推理的“符号派”；受人脑启发搭建结构的“联结派”；受达尔文进化论启发的“进化派”；采用概率推理的“贝叶斯派”以及根据以前出现的类似情况进行推理的“类推派”。

由于人工智能的问题解决方案以及在人类评估算法性能和精确度的人工智能基础的多样性，导致很难对人工智能进行清晰界定，分辨哪些是人工智能构成要素。人工智能研究人员研发了很多用于分析海量数据的技术，这些技术现在被定义为大数据算法和系统。某些情况下，观点可能发生转变。如果某个问题在解决前被认为需要借助人工智能，一旦解决方案被人所熟知，这个问题便会被视为例行数据处理。虽然人工智能的界限并不确定并且随时间

推移发生变化，但这并非问题的重点，重点在于人工智能的研究和应用始终围绕自动化或者复制智能行为。

## 人工智能的现状

在所谓的“弱人工智能”研究方面，科学家已经取得显著进步。这种人工智能针对的是特定应用领域，例如策略弈棋、语言翻译、自动驾驶汽车和图像识别。弱人工智能可以改善很多商业服务，例如旅行规划、客户推荐系统和定向广告。在医疗诊断、教育和科学研究领域，弱人工智能也有重要应用。所有这些都能创造很大的社会效益，增强美国的经济活力。

强人工智能与弱人工智能相对，有时也被称之为“通用人工智能”。这里的“强”指的是一个未来的人工智能系统概念。届时，人工智能系统表现出的智能行为至少达到与人类一样的程度，能够完成各种认知任务。强人工智能的研究面临更为艰难的挑战，成为当前的弱人工智能难以逾越的一道鸿沟。在长达几十年的研究中，科学家一直试图通过扩展弱人工智能解决方案的方式实现强人工智能，但始终鲜有进展。现在，私营部门的专家普遍认为强人工智能至少在几十年内无法实现。美国国家科技委员会也持这种观点。

长久以来，人们一直在思考计算机的智能超越人类将带来何种影响。有人预测称，拥有足够智慧的人工智能可以承担起研发更先进、更智能系统的任务，这些系统随后接替它们的工作，研发更先进的人工智能。理论上，这最终会引发“智能大爆炸”或者达到所谓的“奇点”，让机器的智能迅速超越人类。

在这一进化过程的反乌托邦版本中，超智能机器超越人类的理解能力或者摆脱人类控制。如果计算机掌控大量关键系统，后果将是灾难性的。最好的情况下，人类将无法控制自己的命运。如果是最坏情况，人类将因此走向灭绝。这种设定是科幻小说的惯用桥段。最近，一些颇具影响力的行业领导者纷纷站出来，表达了对人工智能未来发展的恐慌。不过，很多研究人员对未来有着更积极的看法。在他们看来，不断进步的智能系统能够充当人类的助手、助理、培训师和队友。在设计上，它们能够安全运转，同时遵守伦理道德。

美国国家科技委员会的技术委员会的评估结果是，对超智商强人工智能的长期担忧几乎不对当前政策产生任何影响。如果这种恐慌被证明合理，联邦政府应在短期到中期内采用现行政策；如果无法证明其合理性，联邦政府会采取几乎相同的政策。对于长期存在的猜测性风险，最理想的解决之道就是提高能力，降低甚至消除当前已经出现的非极端性风险，例如现在的安全和隐私风险，同时加大研究投入，提高长期应对挑战的能力。此外，随着该领域的研究和应用继续走向成熟，政府和商界的人工智能实践者应适当思考人工智能发展所能带来的长期社会和伦理问题，而不仅仅是技术问题。虽然这种谨慎做法会让一些人相信邪恶的超级智能或可在将来的某一天成为可能，但这种恐慌不应成为人工智能公共政策的主要影响因素。

## 机器学习

机器学习是人工智能最重要的技术手段之一，同时也是人工智能近期取得的很多进展和商业应用的基础。现代机器学习是一个始于大量数据的统计学过程，试图通过数据分析推导

出规则或者流程，用于解释数据或者预测未来数据。这种从数据中学习的方式与过去的“专家系统”人工智能研究方式形成了鲜明对比。在后一种研究途径中，程序员与人类专家协作，学习用于做出决策的规则和标准，随后将这些规则转化为软件代码。专家系统旨在模仿人类专家遵循的规则；而机器学习则依靠统计学方式自行寻找能够在实践中发挥功效的决策流程。

机器学习的优势在于，可以在无法或者难以制定显式规则的情况下采用，最终解决问题。例如，一家经营某项在线服务的公司可以利用机器学习，探测欺诈性用户登录尝试。这家公司可以从以往登录尝试的大数据集着手，凭着后见之明将每次登陆尝试标记为欺诈或者非欺诈。基于这个数据集，这家公司可以利用机器学习推导出用于未来登录尝试的规则，预测哪些尝试更有可能是欺诈行为，应采取额外的安全措施。从某种程度上说，机器学习并非一个用来解决特定问题的算法，而是一种更加通用的解决方案寻找手段，能够利用相关数据解决大量问题。

为了应用机器学习，实践者可以从历史数据集着手，将其分为一个训练集，一个测试集。随后，实践者选择一种模型或者数学结构，利用可调整的参数揭示一系列可能决策规则的特征。这个模型通常被比作一个采用某种规则的“盒子”，参数则是盒子前面可以调整的旋钮，用于控制盒子。实践中，一个模型可能有数百万个参数。此外，实践者还可以定义一个目标函数，用于评估特定参数得出的结果的满意度。这个目标函数通常包含两个奖励要素，一个奖励与训练集较为匹配的模式，另一个奖励更简单规则的使用。

对这个模型的训练过程就是不断调整参数，以优化目标函数。在机器学习领域，训练是一个技术难点。一个含有数百万参数的模型产生的可能结果是一个天文数字，超过你希望尝试的任何算法。如果想让训练算法取得巨大成功，这个算法在探索参数设置空间时必须表现出足够的智慧，能够在具有可行性的运算时间内发现非常出色的参数设置。

一旦模型完成训练，实践者可以利用测试集评估模型的准确性和效用。机器学习的目标是训练出具备归纳概括能力的模型。也就是说，模型不仅在分析训练集的例子时表现出很高的精确度，对未来遭遇的前所未有的情况也是如此。虽然很多模型在完成图像标记等狭隘任务时的表现超过人类，但即便是最出色的模型，也会以无法预知的方式出现失误。例如，对于很多图像标记模型来说，它们可能创建出这样的图像，在人眼看来显然是随机噪点，但经过训练的模型却错误地标记为一个可信度极高的特定物体。

机器学习应用的另一个挑战是，通常无法给出一个直截了当的解释，阐述一个经过训练的特定模型为何在实际运用中起作用。由于经过训练的模型拥有大量可调整参数——通常可达到数亿或者更多——训练虽可以获得一个与数据匹配的有效模型，但并不一定是最简单有效模型。在人类的决策过程中，任何模糊环节通常都归咎于缺少有关决策原因的足够信息，因为决策者可能无法清晰解释为何觉得某项决定是正确的。借助于机器学习，与决策过程有关的一切都成为已知，并且拥有数学保障的精确度，但由于信息太多，可能无法给出清晰解释。

## 深度学习

最近几年，深度学习这个机器学习的分支取得了一些最引人注目的成就。深度学习也被称之为“深度网络学习”，利用启发自人脑的结构。这些结构由大量类似神经元的组件构成。每个组件结合一组输入值，产生一个输出值，随后进入下游的其他组件。例如，在一个图像识别应用中，第一层组件可能结合图像的原始数据，识别图像中的简单图案，第二层结合第一层的结果，识别图案中的图案，第三层结合第二层的数据，以此类推。

深度学习网络通常采用很多层——有时超过 100 个——每层利用大量组件，能够识别极度复杂的数据样式并且精确度极高。最近几年，随着更大更快的计算机系统的出现——让使用更大的深度学习网络成为一种可能——有关如何构建和训练深度网络的新理论也随之产生。超大型网络在执行很多机器学习任务时取得巨大成功，一些专家对此吃惊不已。这也是当前的人工智能研究人员和实践者掀起机器学习浪潮的主要原因。

## 自主和自动化

人工智能通常应用于可以控制物理驱动器或者触发在线活动的系统。当人工智能与日常世界发生联系，自主、自动化和人机合作的问题便会出现。自主是指系统在少量或者没有人类控制的情况下，自行运转并根据环境变化做出调整的能力。例如，一辆自动驾驶汽车可以自行驶往目的地。尽管当前的焦点是汽车和飞机，自主技术实际上是一个范围很广的概念，包括自动交融交易和自动内容策展系统等应用。自主系统还包括可在运转过程中自行诊断并修复故障的系统，例如发现和修复安全漏洞。

自动化是指由机器完成此前由人类负责的工作。自动化涉及到可以被人工智能取代的体力劳动、脑力劳动或者认知工作。至少从工业革命以来，自动化及其对就业产生的影响就是一个重要的社会和经济现象。人们普遍认为人工智能会让某些工作实现自动化，但这是否就是自动化历史的下一篇章，现在还存在很多争论。人工智能能否像过去的自动化浪潮一样对经济产生深远影响？人们对这个问题也存在分歧。

## 人机合作

自动化的目的是让机器取代人类的工作，与之形成对照的是，在某些情况下，需要机器辅助人类工作。这或许是人工智能发展的一个附带结果，又或许是带着打造人机合作团队的目的，而专门研发的系统。用于辅助人类认知能力的系统有时也被称之为“智能增强”。

在很多应用中，人机合作的效率超过其中任何一方。这种方式能够利用一方的优势，弥补另一方的劣势。国际象棋就是一个典型例子。如果给一个实力较弱的计算机选手配备人类队友，这个选手经常能够击败实力超过自己的计算机。即便是实力远超过任何人类的顶级计算机，这样的结果依然成立。另一个例子是 X 光片诊断。在最近进行的一项研究中，参与者要对淋巴结细胞的图像进行诊断分析，判断细胞是否感染癌症。基于人工智能的分析方式的错误率为 7.5%，人类病理学家的错误率为 3.5%，而人机合作的错误率只有 0.5%，降幅高达 85%。

## 对外宣传与报告撰写

该报告由美国国家科技委员会的机器学习与人工智能分委员会完成，该委员会于 2016 年 5 月成立，负责跨部门协调工作，就人工智能相关问题提供技术和政策建议，监督各行业、研究机构及联邦政府的人工智能技术研发。美国国家科技委员会的科技委员会对报告进行了评估，认可报告内容。2016 年，白宫科技政策办公室牵头组织了一系列公共宣传活动，包括与报告中提及的大学和其他协会联合举办的 5 场公共研讨会。该报告在这些活动之后完成。

这五次公共研讨会的具体情况如下：

- 人工智能，法律和治理（5 月 24 日，西雅图，由白宫科技政策办公室、国家经济委员会和华盛顿大学共同主办）；
- 为社会造福的人工智能（6 月 7 日，华盛顿，由白宫科技政策办公室、美国人工智能协会和计算社区联盟共同主办）；
- 人工智能的未来：在全球创业峰会的新兴话题和社会福利（6 月 23 日，帕洛阿尔托，由白宫科技政策办公室和斯坦福大学共同主办）；
- 人工智能的技术、安全及控制（6 月 28 日，匹兹堡，由白宫科技政策办公室和卡内基梅隆大学共同主办）；
- 人工智能的社会和经济影响（7 月 7 日，纽约，由白宫科技政策办公室，美国国家经济委员会和纽约大学共同主办）。

在上述每一场研讨会中，私营部门主办方都举行了有政府工作人员出席的单独会议。超过 2000 人亲自出席了公共活动，此外，在国际互联网上，也有网友参加了活动，仅华盛顿的一场研讨会就有 3500 多名网友参加。

白宫科技政策办公室还发布了信息请求书，以征集公众对这些研讨会话题的评价。信息请求活动截止于 2016 年 7 月 22 日，共收到 161 份反馈。2016 年 9 月 6 日，白宫科技政策办公室对信息请求的反馈意见进行了公布。

## 公共物品人工智能应用

人工智能和机器学习之所以前景极为乐观，原因之一就在于它们具备巨大的潜力，可以解决全世界最困难的挑战及效率低下等问题，从而让人们过上更好的生活。有人将人工智能的前景与先进移动计算技术带来的变革性影响相提并论。在医疗卫生、交通、环境、司法、以及经济包容性等诸多领域，在基础和应用研发方面所进行的公共投资及私人投资，早已让人们开始收获丰厚的回报。

在沃尔特里德医学中心，退伍军人事务部正在利用人工智能，以提高对医疗并发症的预测水平，并改善严重战争创伤的治疗方法，起到了良好的效果：愈合更快，费用更低。同样的通用方法也在——通过对并发症进行预测，医生得以采取预防性治疗——在约翰·霍普金斯大学医院里得到的应用，病人在医院的感染几率也得到了减少。目前，医疗病历正在向电子化方向转变，在诸如精细医学和癌症研究等许多领域，对健康数据进行预测分析起到了至关重要的作用。

在交通方面，由人工智能支持的智能交通管理应用也在发挥作用。在部分地区，人工智能减少的等待时间、能源消耗和尾气排放高达 25%。有不少城市都开始利用叫车服务来进行派车及路线规划，并将其并入公共交通调度和跟踪软件之中，这样一来，公共交通服务就可以变得更及时，更快捷，更廉价，而且在大多数情况下都面向更广大的群众。

有一些研究人员正在利用人工智能技术，通过人工智能图像分类软件对游客们发在社交网站上的照片进行分析，进而更好地对动物迁徙活动进行跟踪。这种软件可以从照片当中识别单只动物，并利用照片的数据及位置标记，建立动物迁徙活动数据库。在白宫科技政策办公室举办的“为社会造福的人工智能”研讨会上，研究人员们对鲸鱼和非洲大型动物的数量及迁徙情况展开了讨论，希望为此而建立迄今为止最大的可用数据集。此外，他们还讨论了海龟的迁徙，希望启动“海龟互联网”项目来跟踪海龟活动，从一个崭新的视角来研究海洋生物。还有发言人谈到，可以利用人工智能帮助反偷猎人员优化巡逻策略，并设计栖息地保护战略，最大程度地保护濒危动物生物多样性。

目前，自动帆船和自动船只早就在海洋上进行巡逻了，所配备的精密传感仪收集着各种数据，包括北极冰川的变化、以及监测海域多变的生态系统数据，对于载人船只来说，这些既昂贵又危险。自动船只与载人船只相比，操作成本要低得多，而且将来还有可能会用于更多领域，比如提高天气预报准确度、加强天气监测、监控非法捕鱼行为等。

在刑事司法方面，人工智能同样具备改进许多工作的潜力，比如犯罪报告、警务工作、保释、量刑及假释决定等。政府正在进行研究探索如何利用人工智能，对目前的做法进行完善，例如“数据驱动正义计划”、及“警察数据计划”等，这些数据计划旨在为执法者和公众提供数据，更好得昭示刑事司法制度中的判刑过程，同时，也要留意因可用数据不足而引发的、人工智能可能产生的偏差和不准确。

部分美国学术机构已经多措并举，利用人工智能来解决经济与社会问题。例如，芝加哥大学提供了一个学术项目，研究利用数据科学和人工智能来解决诸如失业及辍学等社会难题。南加州大学成立了社会人工智能中心，致力于研究如何通过计算机博弈理论、机器学习、自动规划及多主体推理技术等多种学科，来解决无家可归等社会问题。同时，斯坦福大学的研究人员还试图利用机器学习技术解决全球贫困问题。他们利用人工智能对卫星图像中显示的、可能的贫困地区进行分析，以此确定哪里最迫切需要帮助。

许多应用都通过人工智能技术为公众带来了好处，它们依靠的都是可用于训练机器学习模型、并测试人工智能系统性能的数据。如果所提供的数据不涉及个人隐私或商业机密的话，

那么相关的机构和组织就可以向研究人员提供这些数据，帮助他们实现人工智能的继续发展。有了标准化的数据模式和格式，制作可用新数据集的成本和难度也会有所降低。

**建议措施 1:** 鼓励私人 and 公共机构进行自我审视，判断自身是否能够、且通过何种方式，才能尽可能得利用人工智能和机器学习技术为社会造福。社会争议机构和公共政策机构在日常工作中通常不会使用先进技术及数据科技，但应该考虑与人工智能研究人员及从业者开展合作，促进人工智能技术解决更广泛的社会问题，尤其是这些机构正在通过其他方式进行解决的那些社会问题。

**建议措施 2:** 联邦机构应该将人工智能的公开培训数据及公开数据标准放在工作中的首要位置。政府应该强调发布数据集的重要性，这样，人工智能就可以用于解决社会难题了。潜在步骤可能包括：通过实施“人工智能公开数据”计划，实现公布大量政府数据集的目标，加速人工智能研究，在政府、学术机构和私营部门等领域，促进人工智能公开数据标准的使用和最佳实践。

## 人工智能在联邦政府的应用

联邦政府正在研究相关政策和内部实践，以期将人工智能的经济和社会效应发挥到最大，并推动其创新。这些政策和实践有可能包括以下内容：

- 对基础和应用研发项目进行投资；
- 成为人工智能技术及其应用的早期客户；
- 支持试点项目，根据真实情况搭建测试平台；
- 向公众提供数据集；
- 提供赞助支持；
- 确定并追求“大挑战”，为人工智能树立具有可行性的远大目标；
- 为人工智能应用的严格评估提供经费，考察影响及成本效益；
- 为创新的蓬勃发展提供政策、法律和监管环境，同时保护公众免受伤害。

## 在政府中使用人工智能技术，提高服务水平，造福美国人民

在利用人工智能提高服务水平的过程中，存在一个难题，那就是联邦政府提交人通过促进和利用创新来提高为人民服务的水平。但是针对不同的机构，联邦政府开展工作的能力差

异很大。有些机构更侧重于创新，尤其是那些研发经费较多，员工中包括许多科学家和工程师，拥有勇于创新尝试的文化且与创新私营企业开展紧密合作的机构。其中，有许多机构还下设了专门进行高风险、高回报项目的组织（例如，国防部、能源部以及情报部门的高级研究项目机构），从基础研究到前沿发展，对各个领域提供全方位研发经费支持。还有其他机构则将研发作为其主要任务，比如美国国家科学基金会。

但是，有一些机构在相关能力、资源和专业技术等各方面都比较一般，尤其是那些负责减少平困、增加经济和社会流动性的机构。比方说，美国国立卫生研究院的研究预算超过 300 亿美元，而劳动部的研发预算仅为 140 亿美元。这限制了劳动部探索各种人工智能应用的能力，比如利用基于人工智能的“数字教官”技术，加强没有接受过高等教育的劳动者的技能，并提高他们的收入。

国防高级研究计划局的“教育优势”计划，正是发挥人工智能潜力、实现并加速完成重点工作的典范。该计划利用人工智能，能够让新手与数字教官模拟互动，旨在将海军新兵培训为技术专家的时间从几年缩短到数月。一份对数字教官计划的评估报告做了这样的总结：接受了数字教官训练的新兵，在笔试和实战测试中，往往能将服役了 7-10 年的海军老兵给比下去。

基于数字教官试点工程的初步证据还表明，完成了数字教官训练的人员更有可能找到高科技岗位的工作，从而大幅提高他们的收入。这种工资的上涨要远远超过当前的劳动力发展计划所带来的收入增加。在理想状况下，这些结果有可能通过独立进行的、随机的、有控制的试点实验来进行确立。目前，开发数字教官的成本很高，而在开发有效数字教官方面，也没有可以重复进行的方法。通过研究，类似数字教官这样的新兴行业会不断涌现，这些研究对劳动者掌握紧缺技能将会起到潜在帮助作用。

**建议措施 3：**联邦政府应该探索新方法，提高关键机构应用人工智能完成任务的能力。例如，联邦机构应该发挥潜能，设立类似国防高级研究计划局这样的机构，以支持高风险、高回报的人工智能研究及其应用。教育部就是一例，该部提出建立专项高级教育研究项目，以确定人工智能及其他技术是否能够显著提高学生的学习成绩。

**建议措施 4：**美国国家科技委员会的机器学习和人工智能分会应该为政府的人工智能从业者们开辟一个实践社区。各机构应该共同努力，制定并分享人工智能在政府工作中的使用标准和最佳实践经验。各机构应该确保将人工智能引入到“联邦雇员培训计划”的可能性。

## 人工智能与监管

人工智能广泛应用于许多产品中，例如汽车和飞行器。这些应用必须符合相关规定的要求，以保护公众免受伤害并且保证经济竞争的公平性。那么，产品使用了人工智能会怎样影

响相关的管理办法呢？总体来说，管理人工智能产品和保护公共安全应当明晰两个方面：评估人工智能可能减少的风险、以及它可能增添的风险。另外，如果风险处于现有的管理制度范围内，就应该从政策的角度讨论清楚现有的规范能否充分应对风险，是否需要为适应人工智能调整。此外，如果人工智能的管理响应提升了遵守规范的成本，或者是减缓了有益创新的发展或采用，那么决策者就应该考虑调整这些管理办法以降低成本、去除创新的障碍，同时不会危害公共安全或者市场公平。

信息邀请书的评论者大多认为：现在，不宜对人工智能的研究或者应用进行大规模的管理。他们认为，当前的管理目标和管理结构足以应对需求，并呼吁调整现有管理措施以应对人工智能的影响。例如，评论者们建议应当改进机动车辆的管理，以应对自动驾驶车辆的到来。这种必要的改变可以在当前的车辆安全管理架构内进行。在此过程中，相关机构必须牢记管理的根本目的和目标——保护公共安全，同时为人工智能的创新和发展创造空间。

对人工智能等科技的有效管理，政府机构自身需要拥有技术专家来帮助指导管理决策。不论是管理部门和管理机构，还是管理流程的所有阶段，都需要高级技术专家的参与。我们可以利用一系列的人员任命和交流模式（如招聘部门）来培养联邦的工作人员，以获取对当前技术发展状况更多样的视角。政府间人员交流项目等机构就是一例，该项目会在以下的机构间临时雇佣和调配人员：联邦政府、州政府和地方政府，学院和大学、印第安人部落政府、联邦资助的研发中心及其他有资质的组织。如果使用得当，该项目可以帮助各机构满足难以填补的岗位招聘需要，雇佣多种技术背景的人员。项目内的联邦雇员可以承担招聘和联络工作。例如，有的机构人员被派往学院和大学做指导员，鼓励学生加入联邦政府。同样的，有的项目雇员会在不同的工作和领域内轮转，这能够帮助政府雇员获得知识和经验以了解各项规章制度，尤其是人工智能等新兴技术。

**建议措施 5：**在制定人工智能驱动的产品政策时，相关机构应当吸收适当的高级技术专家。有效管理人工智能驱动的产品需要以下三方的合作：机构领导、了解现有管理框架和办法的雇员、人工智能技术专家。机构领导应当采取措施外聘必要的技术专家，或者从内部挑选专家，并且确保在管理政策讨论中保留足够的技术“讨论席位”。

**建议措施 6：**机构应当充分使用人员任命和交流模式（如招聘部门），以培养联邦的工作人员，从多个角度了解当前的技术发展状况。

### 案例研究：自主车辆与飞行器

目前，有关机构正面临更新法律法规以管理基于人工智能的新产品的挑战，相关的例子可见于美国交通部关于自动车辆和无人飞行器的管理工作。就交通部而言，自动驾驶的车辆归美国国家公路交通安全管理局管理，而飞行器归美国联邦航空管理局管理。

## 自动化的前景

公众对于人工智能在车辆上的应用可谓充满遐想。近期推出的新车已经具有了人工智能辅助功能，例如自动停车、保持车辆在规定路径行驶的先进导航控制系统、基于周围车况调整速度等。现在，由人类司机监控的全自动化试验车辆已经上路。专家一致认为，自动化车辆技术最终会比人类司机更安全，而且某一天将会避免每年在国家公路和高速路上发生的数以万计的交通伤亡。

自动驾驶的车辆还可能提高老年人和不能开车的残疾人的行动能力。对于远离工作单位、医疗设施、商店等必要服务的社区来说，来往接驳和其他新型交通方式将为它们带来前所未有的机遇。精密设计的自动驾驶车辆系统能够预测并避免撞车，这可能会极大地减少排放和能耗。政府正在采取措施使这种设想变成现实，例如美国交通部提议在 2017 总统财年投资 39 亿美元进行自动化和互联车辆的研发和配置，以确保美国保持自动驾驶车辆技术的领先水平。

自从 20 世纪 90 年代早期人类可以飞行以来，商用无人飞行器的发展在美国国家空域系统内部一直受到限制。直到最近，无人飞行器主要支持政府行动，例如军事和边界安全行动。但是在最近几年，潜在的应用已经迅速扩展到其他领域，包括空中摄影、测量土地和农作物、监测森林火灾、灾害响应、检测关键基础设施等。几个政府机构正在使用无人飞行器来改进他们的工作，数以千计的美国人已经从美国联邦航空管理局获得商用无人飞行器的必要权限，美国联邦航空管理局的一些做法也加速了这个过程，例如《小型无人飞行器管理规范》在 2016 年 8 月生效，并且在 2015 年 12 月就启动了小型无人飞行器注册服务。美国联邦航空管理局估计截至 2017 年 8 月，美国商用无人飞行器注册数目可能会超过 60 万。

有人估测，在无人飞行器进军航空业的前三年，其经济价值就有可能超过 136 亿美元，且之后会持续增长。无人飞行系统国际协会在 2013 年的一项研究显示，在接下来的十年里，商用无人机行业可以为美国经济创造 820 亿美元的价值和 10 万个新工作岗位。在无人机进入航空业的前十年，将会为美国政府上缴超过 4820 亿美元的税收。

## 确保安全

政府需要采取一系列举措，确保空域和道路安全，同时继续培养创新和增长文化，如此才能让这些具有发展前景的技术的潜在效益成为现实。美国拥有世界上最安全同时也最复杂的航空系统，公众则依靠联邦航空管理局来制定安全标准。《联邦机动车安全标准》对制造商研制安全地面车辆提出了具体要求。在存在不合理安全风险情况下，国家公路交通安全管理局有权召回车辆。在当前的安全管理规定下，车辆每行驶 1 亿英里允许发生 1 起死亡事故，虽然自动驾驶汽车继续减少交通死亡事故的空间很大，但达到或者超过这一安全表现无疑是极大的挑战。

在对安全如此苛求的环境下，采用人工智能技术将面临一系列挑战。首先，需要将驾驶过程中司机负责的内容转化成软件。与弱人工智能的其他成功应用不同，当前并没有车辆或

者飞行器驾驶任务的简要说明。每一种驾驶都涉及多个层面，责任包括导航、探测和避让障碍物、处理瘪胎等机械故障。虽然导航或者探知确定类型对象等子任务与当前的弱人工智能解决方案别无二致，但如何将这些任何整合起来，并设定优先顺序就完全不同了。遵守所有交通规则听起来很简单，但即便是一名经验丰富的驾驶者也可能穿过双黄线，以躲避事故或者超过并排车辆。虽然这种情况很少见，但不应被忽视。我们只需简单计算便可发现，为了让故障发生频率至少降至与人类驾驶者一样的程度，系统必须在不出现故障的前提下处理大量罕见情况。

对于依赖机器学习的系统，将罕见情况考虑在内将影响到系统的设计和测试。如果训练中类似情况很集中，那正确应对就很容易，机器学习解决方案也无疑会更令人信服。当前的挑战是，如何建立一个含有足够罕见情况的数据集。商业航空业拥有整个行业共享事故和安全数据的机制，但报告数据可能并非最近才获得资格认证的无人飞行器操作者本能会去做的事情。对于传统航空业的安全和问责文化，他们还比较陌生。

当前没有具有可比性的机制，供汽车业参考——只有致命事故才会报告，再有就是在州或者地方层面，用一种完全不相干的方式收集和报告其他交通安全信息。断断续续的事故报告或者缺少未遂事故数据增加了确定系统安全所需的里程数或者驾驶时间，对需要进行大量测试加以验证的确定人工智能解决方案构成障碍。

为了进行安全测试，联邦航空管理局指派了 6 个无人飞行器测试场，同时在场内为无人机操作人员提供全面授权。测试场内进行的活动包括，扩展国家航空航天局针对无人飞行器交通管理进行的多年研究，用以确定在低空进行无人飞行器大规模超视距操作的要求。类似地，车联网试点项目以及在俄亥俄州哥伦比亚——2016 年交通部 4000 万美元智能城市挑战赛的得主——部署自动驾驶车辆等地面车辆测试，也将为人工智能研究人员提供丰富的基线和交互数据。

**建议措施 7：交通部应该与产业界和研究人员合作，加强用于安全、研究和其他目的的数据共享。未来，人工智能将在地面车辆和飞行器上得到重要应用。面对这样的未来，联邦机构应在短期内建立起丰富的数据集——收集数据时要注意保护消费者隐私——以便在这些技术成熟时能够做出更合理的决策。**

### 调整现有监管措施

虽然美国国家层面对空域和高速公路的监管路径有所不同，但在整合自动机动车和飞行器监管措施这一点上，双方的目标是一致的：美国联邦航空管理局和美国国家公路交通安全管理局都在努力建设既能确保安全，又能鼓励创新的灵活监管框架。

在空域监管方面，将无人飞行器安全加入空域管理的重大一步，就是 2016 年 8 月 29 日美国联邦航空管理局颁布第 107 款的最后一条，即“小型无人飞行器”规定。该条规定要求现在用途广泛的非娱乐用无人飞行器的重量应低于 55 磅。该条规定要求无人飞行器只能在白天飞行，高度低于 400 英尺，由有资质的操作者控制，且不得脱离操作者直视范围。从

人头顶飞过也被禁止。相关部门计划发布后续规定，随着时间推移和数据累积让人们了解到如何才能安全使用无人飞行器，上述限制会被逐渐放宽。尤其是美国交通部正在制定一项“关于规则制定提议的通知”，提出应当建立一套机制，允许某些“小型无人飞行器”从人头顶飞过，预计之后会进一步扩展允许操作的范围。

美国联邦航空管理局尚未公开发布通往允许全自动飞行的清晰规章路线图。虽然将自动飞行器完全整合到空域管理将经历一个复杂的过程，美国联邦航空管理局正在为不远的将来做出准备，到那时，有人和无人飞行器将在一体化空域体系下共同翱翔。

新的空域管理路径还包括利用 AI 来改进航空交通控制系统。由于在地空实践中使用空中/地面一体化的限制以及对人-人之间沟通的依赖，现有的空域管理体系很可能无法适应未来空中交通运行的密度和多样性。2007 年，也是我们拥有可靠的公共数据的最近一年，由于空中交通延误给美国造成的损失大约为 3100 万美元——从该年起，乘客人数一直在不断上升，因此该数字也会相应增加。虽然受天气和其他因素所限，一些航班的晚点不可避免，但采用新的航空技术、实施新的政策、对基础设施进行升级换代都将大大地提高美国空域的运行效率。这些方案中就包括使用人工智能和机器学习体系，对广大空域使用者，包括有人驾驶和无人驾驶的飞行器，进行更好的安排，在确保安全性的情况下，实现空域的更高效运行。开发和使用这些技术有助于提高空域使用者和服务提供商的全球竞争力，并增加安全性，降低成本。

至于地面交通，目前最为重要的一步就是用于建立常规体系的联邦自动化车辆政策，该政策由行政机构于 2016 年 9 月 20 日正式发布。该政策包括几个部分：

- 为制造商、开发商和其他组织提供的指引，列出了 15 条关于高度自动化车辆的安全设计、开发、测试和应用的“安全评估指南”；
- 州政策模板，对联邦和各州的责任进行了明确划分，为各州政策的制定提供了建议，目标是为自动化车辆的测试和运行建立连贯性的国家制度框架，并为各州的实验留下空间。
- 对目前的制度工具的分析，美国国家公路交通安全管理局可协助自动化车辆的安全开发，如对目前的规则进行解读，允许设计中存在一定的灵活性，但对非传统车辆设计的测试仅允许有限的免检，确保道路上不会出现不安全的自动化车辆；
- 对新的工具和机制的讨论，相关机构未来可考虑支持新的救生技术的安全有效应用，确保道路上使用的相关技术都安全可靠。

随着新数据的获取和新研究的完成，美国交通部会定期对该指南和政策模板进行更新。

**建议措施 8：**美国政府应投资开发和应用一种高端的自动空中交通管理系统，该系统可扩展升级，也能够同时应对无人机和有人驾驶的飞行器。

**建议措施 9:** 美国交通部应继续致力于建立不断完善的制度框架，确保全自动车辆和无人飞行器的安全整合，包括将新型车辆设计融入到公共交通系统。

## 研发与从业者培养

通过在技术研发领域进行投资，培养出一批具有熟练且多样技能的从业者，并掌控不断发展的 AI 技术在经济领域所产生的影响，政府也对人工智能发展发挥着重要作用。本报告发布的同时，一份名为《国家人工智能研发战略计划》的文件也同时发表。这一章讨论的是针对人工智能技术研发及从业者所专门制定的相关政府政策。

## 对人工智能进展的监控

考虑到人工智能所能带来的潜在影响，如果能够准确及时地监控和预测人工智能技术的发展，社会将会从中收益。已经有多个项目试图预测人工智能的未来。2009 年美国人工智能协会关于人工智能长远未来的主席讨论，以及 2015 年人工智能未来研讨会让全世界的人工智能专家汇聚一堂，预测各自领域未来的发展方向。除此以外，斯坦福大学的“人工智能百年研究项目”计划“针对人工智能在自动化、国家安全、心理学、道德、法律、隐私、民主以及其他问题上所能产生的影响，定期开展一系列的研究。”该项目的第一份研究报告已经于 2016 年 9 月发表。

一个极具潜力的研究方法就是长期调查并记录专家们在这一领域的看法。例如，一项关于人工智能研究人员的调查报告显示，百分之八十的受访者相信，和人类智商同级别的“强人工智能”能够最终实现，而其中一半的人相信在 2040 年之前实现这一目标的概率超过 50%。绝大多数受访者相信强人工智能将逐渐超越普通人类的智力。就如上面所讨论的，虽然这些具体的预测还有很大的不确定性，但调查记录这些专业判断却是十分有用的，尤其当这些判断反复出现，长期下来我们就可以看到这个领域专家思潮的演变。了解共识的一个方法就是举办“预测大赛”，例如预测市场状况，参与者如果能够精准预测的话，就可以获得一定的物质奖励。另外还有研究发现，通过研究分析相关出版物和专利数据，可以准确预测人工智能技术的发展。

目前，学术机构和商业组织所进行的绝大多数人工智能基础研究，都会定期公布自己的研究成果，并在相关学术期刊上发表。如果竞争迫使商业机构对自己的研究成果越来越保密，那么对于研究进展的监控也会越来越难，公众的担忧可能也会随之增加。

一个尤其有价值的研究方法，就是发现那些能够代表或者预测人工智能技术重大突破的里程碑。在多个研讨会和会议上，当我们询问参会研究者，政府如何才能发现人工智能领域发展的里程碑，尤其是预示着强人工智能即将到来的里程碑时，他们给出了三种迥然不同却又紧密相关的答案：

1. 能够成功执行结构更简单、内容更广泛的任务：这一观点认为，从目前的弱人工智能逐渐向强人工智能的转变，也将会逐渐扩展弱人工智能系统的功能，使得一个单

一系统可以覆盖更广但结构更简单的任务。这一方面的一个代表里程碑就是扫地机器人能够像人类一样完成一整套家庭日常保洁工作。

2. 不同“风格”人工智能程序的统一：这一观点认为，人工智能目前依靠的是一套各自独立的程序和指令，以完成相应特定的任务。实现强人工智能必然要将这些独立的程序深度整合。这种里程碑式的发展要求找到一个可以处理广泛应用指令的程序，从而取代之前的多个程序。
3. 解决具体的技术挑战，例如迁移学习：这一观点认为，实现强人工智能的关键不在于程序覆盖范围的快速扩展，也不在于对现有程序的整合，而是在于解决技术中正面临的大的挑战，开拓向前发展的新路。现在人们最常提及的一项挑战就是迁移学习，这一技术旨在创造一种机器学习算法，其结果可以被广泛应用（或迁移）到一系列新的应用程序中。例如，迁移学习可以使一个模型具备将英语翻译成西班牙语的能力，这一算法生成的模型可以将这种能力“迁移”到类似的任务中去，例如将中文翻译成法语或者用俄语写诗等，这样完成这些新任务的学习速度就会快得多。

**建议措施 10：**美国国家科技委员会机器学习和人工智能分委员会应该监控人工智能技术的发展，并定期向上级主管部门领导汇报人工智能的发展状况，尤其是关于其发展技术中的里程碑式突破。随着技术的不断发展和专家共识的不断变化，分委员会应及时更新里程碑清单，并考虑在合适的时机向公众报告人工智能技术的发展状况。

**建议措施 11：**政府应该监控其他国家的人工智能发展状况，尤其是人工智能技术的里程碑式突破。

**建议措施 12：**人工智能相关行业应该和政府合作，及时向政府反馈行业的最新发展动态，包括近期可能在技术上取得的重大突破。

### 联邦政府对人工智能研究的支持

2015 年，据公开数据显示，美国政府在人工智能相关领域投入研发资金大约为 11 亿美元，此前预测显示 2016 年相关投入将增长到 12 亿美元。在白宫科技政策办公室主办的所有人工智能相关研讨会和公共推广活动中，无论是业界领袖、技术专家还是经济学家，都向政府官员呼吁加大在人工智能技术研发方面的政府投入。

人工智能研究领域的领军人物对保持目前相关技术的快速发展，并将其应用于更广泛的领域表现出十分乐观的态度。他们同时强调，现在有许多重大问题还没有找到解决方案，通往强人工智能的道路也还不明朗。

研究人员报告指出，在过去几十年间，对人工智能领域的研究热情和相关投资有着很大的波动，其中一个低潮期被称作“人工智能的寒冬”。同时他们强调，从历史经验来看，一

项计算机技术的重大突破，想完成从实验室中的创意到成熟的工业化生产应用，往往需要15年甚至更长的时间，因此对这一领域进行持续稳定的投资显得尤为重要。

我们有充分的理由支持联邦政府加大在人工智能研究方面的资金投入。经济顾问委员会的分析显示，不仅仅是人工智能研发领域，在所有科研领域，增加两倍乃至三倍的研发投入所带来的经济增长，对一个国家来说也是一项值得投资的净收益。尽管从财政的角度来看，在所有科研领域如此大规模地增加投资是不可能完成的任务，但有针对性地增加对经济回报率高和有战略价值的领域的投资则是完全可行的，这样可以在获得可观收益的同时，避免产生过大的财政负担。人工智能就属于这种领域，研究机构的报告显示，如果让人工智能研究组织获得更多的财政支持，能够实现人工智能技术的快速进步，以及相关专业从业人员的快速增长。2016年7月，在纽约一场人工智能研讨会上，经济顾问委员会主席杰森·弗曼发表了一场演讲，他在演讲中说道：“过去十年来，我们在机器人技术、人工智能和其他许多领域都取得了大量的发明创新。但是，在这些领域我们需要用更快的创新速度，来真正带动生产力和生产效率的高速增长，”他还提到自己对人工智能最大的担忧是“我们没有足够的（人工智能技术）。”

可以肯定的是，私营企业将会是人工智能技术发展进步的主要引擎。但从目前的现状来看，在基础研究方面的投资还远远不够，基础研究投入周期长，研究目的纯粹是为了拓展这一领域的科学边界，因此私营企业很难在短期内获得相应的投资回报。虽然基础研究可以让所有人受益，但是其研究成本却只有负责的企业独自承担。相关的研究报告显示，与能带来最高的经济发展水平的研发投入规模相比，目前的研发投入水平大约只有其规模的四分之一到一半而已。

**建议措施 13：联邦政府应该优先发展人工智能的基础和长期研究项目。如果联邦政府和私营企业能够在人工智能研发领域长期稳定地投入资金，尤其是在高风险的基础研究领域的长期投资，会让整个国家从中受益。由于私营企业尤其不愿意在长期的基础研究领域进行投资，这就更凸显出联邦政府在一领域进行研发投资的重要性。**

### 从业者的发展和多样性

人工智能的快速发展对具有相关技能的人员支持和相关领域的发展也提出了巨大的需求。人工智能从业者包括人工智能研究人员，这些人直接推动了相关技术的进步；大量的技术专家，这些人决定了解决某些特定问题的人工智能程序；数量更加庞大的用户，这些人负责在特定的条件设置下操作这些程序。对研究者来说，人工智能训练要求相关人员有极强的跨学科素质，通常包括在计算机科学、统计学、数理逻辑和信息论方面的深厚积累。对技术专家来说，训练通常强调要有在软件工程和应用程序领域的相关背景。对用户来说，需要熟悉人工智能技术，这样能够熟练应用相关的人工智能技术。

## 政府的角色

人工智能从业者方面的挑战，一定程度上是在科学、技术、工程与数学教育方面的挑战，这也是美国国家科技委员会、白宫科技政策办公室和其他组织关注的重点。美国国家科学基金会和教育部正在和私营企业及各各级政府展开合作，致力于提高教育质量、灵活性以及行业影响，从而突出实现经济可持续性发展、增加包容性和多样性以及改善产出衡量标准方面的目标。美国国家科技委员会科学、技术和数学教育分委员会将众多支持 STEM 教育项目的联邦机构聚集在一起，在多个研究课题上进行协调，其中就包括人工智能教育。

## 学校的角色

将人工智能、数据科学及相关领域与全国教育系统整合起来，对发展出一批突出国家战略重点的从业者是十分必要的。各个级别的教育机构都在设立和发展人工智能项目。大学、院校甚至中学都在扩充人工智能和数据科学课程，但我们现在需要更多的项目和教师。

对这些学术机构来说有几个关键角色：

- 组织和维持研究人员队伍，包括电脑科学家、统计学家、数据库和软件工程师、博物馆和图书馆馆员，以及精通数据学的档案管理员；
- 对相关专业人士进行培训，突出软件开发课程中的人工智能程序，提供可以演示人工智能应用在其他领域中人工智能应用课程，将相关行业、公民社会和政府所面临的人工智能和数据科学方面的挑战，整合成高效案例学习分析；
- 确保操作者对人工智能系统有足够的了解，以满足整个行业、政府和学术机构对用户的要求；
- 通过种子基金、职业发展补贴、实习机会、奖学金及暑期研究实习等手段，给予相关培训方面的支持；
- 招聘并维持相关教职人员，因为对有经验的研究人员来说，行业内的薪资增长速度要超过学术机构的薪资。

社区大学、高等职业院校和学位课程也扮演了非常重要的角色，它们能为学生和职员提供必要的技能培训，同时在时间和金钱方面不用有过多的投入。这些学习机会尤其对于工人扩展个人的职业技能、退伍老兵再就业以及失业人员培训再就业都可能会起到很大的帮助。

在一个具备人工智能技术的社会中，要求全体公民都能够阅读数据、理解数据，对数据进行沟通，并参与到与人工智能相关政策制定的讨论中来。在中学乃至小学就引进数据科学课程，可以有效改善国民的数据知识水平，同时为学生在高中之后理解更高级别的数据科学概念和课程做好准备。

人工智能教育同时还是“全民计算机科学教育”项目的组成部分，总统发起这个项目的目的是，让所有美国学生从幼儿园一直到高中都可以学习计算机科学课程，培养计

计算机思维技能，在数字经济中成为一名创造者而不仅仅是消费者，同时在一个技术驱动的社会中成为一名活跃的合格公民。美国经济正在发生剧变，无论是教育工作者还是商业领袖都越来越意识到，计算机科学在经济机会和社会流动性方面已经成为一门“新基础技能”。在此之前，全国各地的家长、老师、社区、州政府以及私营企业领袖已经在计算机科学教育方面有了很大的投入，全民计算机科学教育项目就是建立在这一基础之上，也是为了让未来的从业者准备好迎接人工智能驱动经济的需求所带来的挑战。

## 多样化挑战

所有领域都在面对人工智能从业者多样化带来的挑战。人工智能从业者没有性别和种族差异，这也反映出技术工业和计算机科学领域也基本没有太多的相关差异。如何充分发掘出美国人民的潜力，尤其是在科学、技术、工程与数学-领域、商业领域和技术工业领域的潜力，是这届美国政府的首要任务。将不同背景、经历和身份的人群吸纳其中，尤其是女性和通常在 STEM 教育中被忽视的种族和人群，是非常重要的，这也是计算机科学和人工智能所面临的最主要的挑战之一。

目前，计算机科学专业的毕业生中只有 18% 是女性，而在 1984 年最顶峰时期这一比例是 37%。尽管在人工智能领域缺少长期稳定的人口统计数据，但我们还是可以找到一些数据。在 2015 年神经信息处理系统进展大会上——这是年度最盛大的人工智能研究研讨会之一——参会人员中只有 13.7% 是女性。Textio 是一家将人工智能应用在工作招聘广告和邮件中的创业公司，在一场机器智能研讨会上，看到同样如此低的女性参会人员比例之后，公司的女 CEO 和创始人决定对这一行业在招聘中使用的语言展开深入调查。公司在对 78,768 份工程师招聘广告进行分析之后发现，机器智能行业在软件工程师招聘广告中的性别偏见得分，是其他任何行业的两倍多。

多样性挑战不仅仅体现在性别方面。相比之下，非裔和西班牙裔美国人以及其他少数民族群在 STEM 领域、计算机科学和技术行业从业者——包括人工智能——占人口中的比例要远远低于各自族群总人数占美国总人口的比例。

白宫科技政策办公室信息邀请书中收到的评论也多有谈到多样性挑战。评论者关注的是人工智能应该是由多样化的人群创造的，也应该是为多样化的人群服务的。这样可以避免过人工智能在过于单一的领域发展所带来的负面影响，包括在开发计算算法时产生偏见的风险，相应的措施就是要更包容，利用拥有多样经历、背景和意见者的力量。在公共研讨会上，人们也广泛深入讨论了这些话题。关于人工智能从业者和人工技能技术设计中缺乏多样性所带来的影响，以及人工智能对社会的影响发面有一些相关研究。这一内容丰富的研究课题有越来越多人参与进来，但依然无法满足更广泛的人工智能从业者发展需求。无疑更多的研究能带给我们更多的收益。

**建议措施 14：**美国国家科技委员会(NSTC)机器学习与人工智能网络信息技术研发分委员会，以及科学技术工程教育委员会应该联合发起一项关于人工智能从业者流水线

的研究，以发展出更多事情及行动，促进从业者包括人工智能研究人员、专家和用户在数量、质量和多样性上的合理增长。

### 人工智能、自动化与经济

短期内，人工智能最突出的经济效益就是可以让过去无法自动化的任务实现自动操作。在新的自动化浪潮中，我们有许多过去的先例可以借鉴，但人工智能在许多方面还有不同之处。政府必须充分了解人工智能的潜在影响，以便制定出相应的政策和制度，发挥人工智能的效益，减少相应的成本。

与过去的革新浪潮一样，人工智能也是效益和成本兼而有之。历史上的自动化浪潮的首要益处就是生产力的提高；如今这一波自动化浪潮也不例外。例如，2015年一项针对十七个国家的机器人研究发现，机器人的使用使这些国家在1993至2007年之间的年度国民生产总值平均增长了0.4个百分点，对这些国家同期的整体国民生产总值的贡献略高于百分之十。

但是，之前的自动化浪潮也引发了一些值得关注的问题，其中包括对某些类型的工作和行业的潜在影响，以及随之而来的收入不平等。由于人工智能有可能会使一些职业从此消失，或者使一些工作薪水降低，尤其是中低技术要求的职业，这时，就需要有政策的干预来确保人工智能的经济效益可以广泛共享，不平等会因此消失而不是加剧。

以人工智能为代表的自动化引发的经济政策问题至关重要，最好交由白宫的独立工作组去解决。白宫应另外组织跨部门的研究，了解自动化对经济的影响，提出应对的政策，并在未来几个月之内公布于众。

**建议措施 15：**总统办公室应在年底前出版一份后续报告，进一步研究人工智能和自动化对美国就业市场的影响，并提出应对政策的纲要。

### 公平、安全与治理

随着人工智能得到越来越广泛的应用，技术专家和政策分析员们对其导致的意外效应也提出了越来越多的关注。使用人工智能为人类做出重要决定，取代了人类本身和相关机构的决定权，并引发了人们对于如何确保公正、公平和责任的关注——这是在“大数据”时代就曾被提出过的问题。使用人工智能控制现实世界的设备还引发了对安全问题的关注，尤其是当这些系统需要面对人类环境的各种复杂境况时。

从技术层面而言，公平和安全方面的挑战紧密相关。针对这两个问题，从业者们尽力防止故意歧视或破坏的产生，避免意外的后果，并向利益相关方提供所需的证据，确保不会发生意外的事故。

## 正义、公平和责任

应将人工智能的社会利益与社会和经济影响作为法律和政府管理的常规课题召开专题讨论会，确保人工智能可以促进正义和公平，确保以人工智能为基础的技术能够取得利益相关方的信赖。这一问题曾在行政部门 2014 年 5 月发布的首份大数据报告和 2016 年 5 月发布的后续报告“大数据、算法系统、机会和人权”中重点探讨过。

在刑事司法系统中，由于数据总量和优质数据不足，大数据饱受质疑。人工智能需要好的数据。如果数据不完整或是有失偏颇，人工智能可能会使问题更加恶化。在刑事司法系统中使用人工智能的任何人都必须对目前数据的有限性给予充分的关注。

在 ProPublica 的一篇深度报告中曾举出了一个例子，这也是在研讨会上常常提到的一个问题，即一些法官在刑事判决和保释聆讯中往往会使用偏差明显的“风险预测”工具，一些监狱管理人员在分配和假释决策中也常常使用。文章中提出的一些证据表明，法官们使用的这种商业风险评分工具会产生种族偏见。Upturn 的另外一份报道也对预测性警务工具的效果和公平性提出了质疑。

同样的问题在招聘中也会遇到。如果使用机器学习模型对职位申请者进行筛选，而该模型中使用的数据反映的是带有偏见的先前决策，那么过去的偏见将会一直延续。例如，要找一位与之前的雇员相类似的求职者，那么系统就会更青睐那些与目前团队成员较为接近的人，而不会在考虑了潜在求职者的各种多样性之后做出最佳的选择。

针对这些问题，一些研讨会的发言者提出，当人工智能应用于公共目的时应该更加公开透明。一位发言人将公共决策中行政机构的角色与人工智能的角色做了对比。出于机构对某些事项的专业性，我们往往会授权给某个机构，但由于流程保护、促进透明度和监督等各种措施的限制，授权的范围非常有限。一些发言人要求开发出类似的理论来规范对机器的决策授权，来确保责任明晰。人们对透明度的关注不仅针对人工智能使用的数据和算法，更关心那些以人工智能为基础的决策是否能由人类进行解读。

在这些研讨会上，人工智能专家们还提出，由于人工智能系统的复杂性及其使用数据的庞大性，我们必须谨慎地了解、预测和解释高级人工智能系统行为本身固有的挑战。

人们往往陷入一个误区，认为复杂的算法一般是遵照其设计者的指令，那么只有在设计者自身有意或无意地拥有偏见时，一种算法才会出现偏见。这一点使得人们难以正确认识机器学习的结果。一名技术开发员若想开发出一种带有偏见的算法自然不是难事，而无意识的偏见也常常会使从业者们防范不足。但是在实践中，无偏见的开发者们尽心尽力开发出的系

统也难免会产生带有偏见的结果，因为即便是人工智能系统的开发者也无法完全了解并防止意外结果的产生。

莫里兹·哈特提出了关于机器学习过程中偏见是如何无意产生的图示。他假设一台机器学习模型可用于区分人们的真名和假名。在数据设置中，如果名字部分比较特殊，那么该模型就会倾向于将其认定为假名。这条规则对于整体人群来说或许有一定的预测能力，因为假名常常不拘一格，所以一般也更独特。但是，如果有一个人数不多的少数民族总是取一些与大多数民族相区别的名字，那么这些与众不同的名字就很可能被认定为独特的样本，并因此被错误地划分为假名。这种结果并不是由于对少数民族名字的特殊处理而造成的，也不是由于输入的数据在整体人群中不具有代表性，只是由于少数民族的人口不够多而已。

卡内基梅隆大学计算机系主任安得烈·穆尔在一次关于人工智能技术、安全和控制的研讨会上就应对人工智能的挑战和未知结果提供了一条思路。他认为，如今，由于人工智能算法的不透明，要减少未知结果的风险，最有效的办法就是进行大量的测试——尤其应该对各种坏的结果列出长长的清单，并通过创建更加有针对性的测试来寻找并清除这些结果。

举个例子，在未进行大量测试时，一个为照片自动添加说明的模型可能会在一些肤色较深的肖像照片上添上“大猩猩”字样。这与开发者的意图背道而驰，即便测试结果显示该模型已经为绝大多数照片提供了准确的结果，可这种糟心事还是会发生。这种类似错误虽然发生的几率很小，但其造成的恶劣结果却不是人力所能左右的，其中并没有蕴含任何种族概念，也没有对相关历史背景的看法。防止这类问题的一个办法就是进行大量的算法测试，对人们的面部进行仔细地检查和标注，包括人为地对一些结果进行筛查，识别出模型无法发现且人类无法接受的结果。

对人工智能的从业者和学生进行道德伦理教育也是该方案的一个重要组成部分。最好能让每个学习人工智能、计算机或数据科学的学生都加入有关道德和安全问题的课程和讨论。但是，单单学习道德还不够。道德可以使从业者们认清他们对所有利益相关方的责任，但道德培训还应增强学生们将良好的意图植入实践的技术能力，在构建和测试系统时，加入更多的技术防范。

由于从业者们致力于使人工智能系统更加公正、公平和易于解释，技术手段也常常成为解读的助力而不是障碍。改进机器学习结果的可解释性的研究就是一个实例。可解读的模型使人们了解决策的流程，并能对其设定的条件和过程提出质疑。

有许多技术手段可用于提供复杂算法决策的可解读性和可靠性，比如可以让系统进行“实地”测试，即为其设定场景，并观察其行为。也可以让系统进行黑匣子测试，只要输入虚拟的数据，并观察其行为，就能测试到非正常情况下发生的行为的情形。系统设计的某些或全部技术细节都可以公开发布，使分析员们能够复制并对其内部行为的方方面面进行分析，而这些都是仅靠测试无法识别的。在有些情况下，公开信息帮助公众对系统的偏见风险进行评估，同时保留系统的其他信息以便保护其专有权或私有权，是完全可以实现的。

## 安全和控制

在研讨会上，人工智能专家们提出，限制人工智能在现实世界应用的一个主要因素就是对于安全和控制的关注。如果从业者无法确保系统的安全性和可控性，确保使用该系统不会产生无法接受的严重的负面风险和结果，那么该系统就不能也不应该得到应用。

安全和控制方面的一个主要挑战就是在建立系统时，必须能够从实验室的“封闭世界”安全地过渡到外面的“开放世界”，而在开放世界中，各种不可预测的事情都可能会发生。在开放世界中，系统可能会遇到其设计和建立时没有考虑到的物体和情形。而有条不紊地适应这些无法预见的情形很难，但对于安全操作却是必不可少的。

提到人工智能的安全和可预测性问题时，一些发言人曾提到最近发表的一篇名为“人工智能安全方面的具体问题”的论文，该文的第一作者曾在研讨会上做了关于技术、安全和控制的发言。该论文使用了一个自主研发的机器人进行家务劳动的实例。在论文的概述部分，作者列出了要确保机器人安全有效，在实际操作中可能引起的各种问题，包括：

**避免负面影响：**我们如何确保清扫机器人在实现它的目标时不会对环境产生负面的影响，如它要快速地清扫却打碎了一个花瓶？我们是否能够人工地输入各种注意事项，确保机器人不会犯错？

**避免对作弊行为的奖赏：**我们如何确保清扫机器人不会利用它的奖赏功能戏弄人类？比如，如果我们因为该机器人把环境打扫得一尘不染而对其进行了奖赏，它会不会关闭自己的镜像功能这样就看不到任何的垃圾了，或者用它无法看透的材料将垃圾覆盖，或者只是当人类在身旁的时候把垃圾藏起来，这样他们就看不出来了。

**可扩展的监督：**我们如何有效地确保清扫机器人能够对昂贵的物件给予足够的重视，并且无需对这些物品进行频繁地估价？例如，它应该把不属于任何人的物品扔掉，但能把属于某人的物品放在一边（它能够把满地乱扔的糖纸和散落在地的手机区别对待）。可以向涉及到的人员进行询问和核实，但这种核实不能太过频繁——机器人是否能够在信息有限的情况下做出正确的选择呢？

**安全的探索：**我们如何确保清扫机器人不会进行探索性的动作，并导致恶性后果呢？例如，机器人可以进行拖地方法的实验，但如果把一条湿拖布放在电源插口处可就大大不妙了。

**分配转变的可靠性：**我们如何确保清扫机器人能够在与其受训环境截然不同的情况下仍能有效识别并可靠地运转？例如，用清理工厂车间的启发式算法来清理办公室可能就是彻彻底底的灾难。

这些例子清楚地表明，人工智能系统的“智能”可能深入却很狭窄：系统可能在发现尘土、优化拖地策略等方面具备超人般的能力，但却不知道避免将湿拖布放在电源插口上。用一句话来形容这种问题：我们如何让智能机器具备常识？研究人员在这类问题上的进展非常缓慢。

## 人工智能安全工程

技术、安全和控制研讨会上一个常见的主题，就是将开放世界的人工智能方法与更为广阔的安全工程领域联系起来的必要性。在建造其他类型的关键安全系统方面的经验，如飞机、发电厂、桥梁、车辆等，对于人工智能的从业者们的验证和确认也非常有借鉴意义，如何实现某项技术的安全性，如何管理风险以及如何与利益相关方就风险问题进行沟通。

目前，人工智能的实践，尤其是在机器学习的飞速发展领域，既是一门学问，也是一门艺术。实践中的一些方面并没有完整的理论支持，而是依赖于从业者们的直觉判断和实验。这在新兴技术领域并不罕见，但却限制了技术在实践中的应用。一些利益相关方希望将人工智能发展成为更加成熟的工程领域。

随着工程领域的日趋成熟，它们就可以摆脱最初的“手艺”阶段。在该阶段，发展的驱动力只是业余爱好者的直觉和动手精神；而要发展到第二个商业阶段，则应包括高技术雇员、务实的改进、广泛接受的经验法则以及有组织的生产和销售；再发展到更为成熟的下一阶段，融合了更加严格的方法、更有技能的专业人才，更为完整的理论体系以及更加专业化的产品。大多数比现代人工智能历史更为悠久的工程领域都已经进入了成熟的阶段。

总之，成熟的工程领域在创造更可预测、更可靠、更强劲、更安全和更有保障的系统方面有着更大的成就。继续这种发展趋势，直到人工智能也成为一个成熟的工程领域，那么即便日后建立更多更复杂的系统，也能保证它们的安全性和可控性了。

**建议措施 16：**使用人工智能系统进行决策或为相关决策提供支持的联邦机构在决定个人事宜时，应当特别注意确保这些系统的效力和公平，并以证据为基础进行验证和确认。

**建议措施 17：**资助各州和地方政府使用人工智能系统进行有关个人的决策的联邦机构应当对资金的使用条件进行审核，确保使用联邦资金购买的以人工智能为基础的产品或服务能够产生足够透明的结果，并以证据确保其效力和公平性。

**建议措施 18：**学校和大学应该将道德、保障、隐私和安全等主题作为人工智能、机器学习、计算机科学和数据科学课程的一部分。

**建议措施 19：**人工智能专业人士、安全专业人士以及他们所在的专业团体应该共同努力将人工智能向着更为成熟的人工智能安全工程领域推进。

## 全球考量与安全议题

除了人工智能长久以来就有的种种挑战和涉及公平安全的具体问题，人工智能对于国际关系、网络安全和国防也构成了重要的政策问题。

## 国际合作

随着国家、多边机构和其他利益相关者开始评估人工智能的好处和挑战，人工智能在近年来的国际讨论中成为了热门话题。这些机构之间的对话和合作有助于推进人工智能的研发并使其发挥益处，同时处理相关挑战。特别是在人工智能方面的一些突破是多个国家的人力、资源和机构合作研究的直接或间接成果。与其他的数字政策一样，各国需要共同努力发现合作机会并制定国际框架，从而促进人工智能研发，应对种种挑战。美国作为人工智能研发的领先国家，可以通过政府间的对话和合作继续在全球研究合作方面扮演重要角色。

国际参与对于全面探索人工智能在卫生保健、制造业自动化和信息通讯技术中的应用必不可少。人工智能应用也可以处理全球问题，例如灾害预防和应对、气候变化、野生动物非法交易、数字鸿沟、就业和智慧城市。美国国务院将隐私问题、无人驾驶的安全性以及人工智能对于长期就业趋势的影响视为需要在国际环境下关注的人工智能政策领域。

为了支持美国外交政策在该领域的优先权——包括确保美国国际领导地位和经济竞争力——美国政府致力于人工智能研发和与其他国家（包括日本、韩国、德国、波兰、英国和意大利以及多边论坛）进行双边讨论的政策问题。国际人工智能政策问题和人工智能的经济影响也向联合国、七大工业国、经济合作与发展组织和亚太经合组织提出。美国政府希望人工智能成为国际活动中日益受关注的话题。

美国致力于与工业和相关标准化组织合作，以促进国际标准朝着工业主导、自愿参与和寻求共识的方向发展，并且以透明、开放和市场需求的原则为基础。美国的措施被正式纳入法律（NTTAA, PL 104-113）和政策（OMB Circular A-119）中，并在美国标准战略中再次重申。

**建议措施 20：**美国政府应在涉及人工智能的国际参与方面制定政府层面的政策，并制作需要国际参与和监督的人工智能热点领域清单。

**建议措施 21：**美国政府应该加深与关键国际利益相关者的合作，包括外国政府、国际组织、业界、学术界等，以交流人工智能研发方面的信息并促进合作。

## 人工智能和网络安全

今天的弱人工智能在网络安全方面有着重要的应用，并且有望在防御（应对）措施和进攻性（主动）措施方面发挥日益重要的角色。

目前，设计和操作安全系统需要专家投入大量的时间和注意力。专业工作自动化——部分或全部——可以使更广范围的系统和应用以明显更低的成本实现较强的安全性，并且增加网络防御的灵活性。使用人工智能有助于维持快速应答，以发现和应对不断变化的网络威胁形势。人工智能（特别是机器学习系统）将有很多机会帮助处理网络空间的复杂性并为有效人类决策提供支持以应对网络攻击。

未来的人工智能系统可以执行预测分析，通过根据大量、不断变化和通常不完整的现有数据来源制定动态威胁模型预测网络攻击。这些数据包括网络节点、链接、设备、架构、协议和网络的拓扑结构和状态。人工智能可能是诠释这些数据的最有效方式，可以主动识别弱点并采取措施预防或减轻未来的攻击。

到目前为止，由美国国防高级研究计划局举办的网络超级挑战赛结果显示了这一措施的潜在发展。网络超级挑战赛的目的是加快先进和自动系统的发展，这一系统可以在对手有机可乘之前，发现、评估和修补软件的安全问题。网络超级挑战赛总决赛在 2016 年 8 月 4 日举行。为了促进后续的研究和平行竞争，网络超级挑战赛总决赛期间公开了由自动化系统产生的所有编码，作为允许其他人进行逆向工程并从中受益的开放资源。

人工智能系统也有自身的网络安全需求。人工智能驱动的应用应该执行稳定的网络安全控制，以确保数据和功能的完整性，保护隐私和机密并且保持可用性。最近公布的联邦网络安全研发战略计划强调了对于“可持续确保系统开发和操作”的需求。网络安全的发展对于人工智能解决方案在应对恶意网络活动时的安全性和恢复力至关重要，特别是在政府和私营部门业务使用弱人工智能的规模和类型不断增加的情况下。

最后，人工智能可以支持计划、协调、整合、同步和指导有效操作和防护美国政府网络和系统的活动，在支持私营部门网络和系统的安全操作方面提供协助并使实际工作符合所有的适用法律、法规和条约。

**建议措施 22：机构的计划和战略应考虑到人工智能和网络安全之间的相互影响。参与人工智能问题的机构应该向美国政府和私营部门网络安全部门确保，人工智能系统和生态系统在应对智能对手时能够保持安全性和恢复力。参与网络安全问题的机构应该使美国政府和私营部门人工智能部门制定创新的方法应用人工智能，以实现有效和高效的网络安全。**

### 武器系统中的人工智能

几十年来，美国一直都在把自动化引入部分武器系统。这些技术改进有助于在使用这些武器系统时实现更精确、更安全、更人道的效果。精密制导武器在作战时将减少武器使用和附带损害，而遥控航空器可以使军事人员和危险之间的距离更大，从而减轻对军事人员造成的危险。但是，如果人类不直接控制武器系统，会有一定的风险，而且会产生法律和伦理问题。进一步结合自动和半自动武器系统与美国国防计划和部队结构的关键是继续确保所有武器系统（包括自动武器系统）的使用符合国际人道主义法。此外，美国政府应继续采取适当的措施控制扩散，与合作伙伴和同盟国共同制定关于武器系统开发和使用的标准。

特别是在过去几年，技术专家、伦理学家和其他国际社区人员提出了关于“致命性自主武器系统”开发的问题。在特定常规武器公约的背景下，美国积极参与持续进行的关于致命

性自主武器系统的国际讨论，并预测针对这些武器系统潜在前进之路的国际讨论将持续稳步进行。

《特定常规武器公约》的缔约国讨论了技术、法律、军事、伦理和其他关于新兴技术的问题，尽管很明显，各国对于致命性自主武器系统未达成共识。一些国家试图将军用无人机纳入致命性自主武器系统的概念。美国反对这一观点，因为根据定义，无人机与载人飞机一样直接由人类控制。其他国家都关注于人工智能、机器人军队，或是“有意义的人类控制”——这个术语尚无良好定义——能否应用于事关生死的决定。美国一直在重申，所有的武器系统，无论自动或非自动，必须坚持国际人道主义法，包括区别原则和比例原则。为此，美国始终认识到了武器审查过程对于新武器系统开发和采用的重要性。《特定常规武器公约》将决定是否以及如何于2016年12月举行的审议会议中讨论致命性自主武器系统和相关问题。

美国政府还对国防系统的自主性的影响进行了全面的审查。2012年11月，国防部颁发了国防部指令3000.09“武器系统的自动化”，概述了开发和使用自动和半自动武器的要求。可自动选择和设定目标并具有致命武器的武器系统在正式开发和使用之前需要高级国防部审查和批准。国防部的指令既不禁止也不鼓励这一开发，但要求谨慎进行，并需经过高级国防官员的审查和批准。除此之外，国防部指令要求自动和半自动的武器系统都经过严格测试，人员在推进关于武装冲突的国际标准方面经过适当训练。

在一系列与国防相关的活动中，人工智能可能提供显著的优势。非致命性的活动，如物流、维护、基本操作、退伍军人的医疗、救生战场医疗救助和伤员运送、人员管理、导航、通信、网络国防和情报分析可以从人工智能中受益，使美国军队更加安全和高效。人工智能也可以通过非致命手段在新的系统中扮演重要角色，保护人员和高价值的固定资产并防止攻击。最后，这些应用可能会成为国防部最重要的部分。

从更广泛的角度看军事技术和人工智能的进步，科学家、战略专家和军事专家一致认为，致命性自主武器系统的未来难以预测，而且变化速度非常快。许多新水准可能不久便会成为达到，并迅速得以发展和运用。行政机构正积极、持续进行跨部门讨论，致力于制定关于自动武器的政府政策，并符合人类共同的价值观、国家安全利益和国内国际义务。

**建议措施 23：美国政府应该在符合国际人道主义法的基础上制定关于自动和半自动武器的统一政府政策。**

## 结论

如果业界、公民社会、政府和公众共同努力，支持技术的发展，密切关注它的发展潜力，管理它的风险，那么人工智能就将成为经济增长和社会进步的主要驱动力。

政府可以发挥多方面的作用。它可以召集各方就重要问题进行交流，帮助制定公开辩论的日程。它可以在技术开发的同时对其应用的安全性和公平性进行监督，并制定规章制度在保护大众的同时鼓励创新。它应该支持基础性研究，支持将人工智能应用于公众福利，支持有技能的、多样化的从业者的培养。同时，政府自身也应使用人工智能，更快捷、更有效、更经济地服务大众。

从教育到经济安全网、再到国防、环境保护、刑事司法，许多公共政策的领域都将在人工智能的持续发展迎来新的机遇和新的挑战。政府应不断提升自身的能力，了解并应对这些改变。

随着人工智能技术的不断发展，从业者们必须确保人工智能为基础的系统能够得到有效的管理；确保它们公开、透明、易于理解；确保它们能够和人们一起高效地工作；确保它们的应用不会与人类的价值观和理想背道而驰。研究人员和从业者们已经增加了对这些挑战的关注，并将继续关注下去。

开发和研究机器智能可以帮助我们更好地理解和鉴别自身的人类智能。缜密周到地使用人工智能可以提高我们的智能水平，帮助我们规划出一条更美好、更智能的道路，大步前进。

## 报告建议措施

为查阅方便起见，本节将报告中的所有建议措施列出如下：

**建议措施 1:** 鼓励私人 and 公共机构进行自我审视，判断自身是否能够、且通过何种方式，才能尽可能得利用人工智能和机器学习技术为社会造福。社会争议机构和公共政策机构在日常工作中通常不会使用先进技术及数据科技，但应该考虑与人工智能研究人员及从业者开展合作，促进人工智能技术解决更广泛的社会问题，尤其是这些机构正在通过其他方式进行解决的那些社会问题。

**建议措施 2:** 联邦机构应该将人工智能的公开培训数据及公开数据标准放在工作中的首要位置。政府应该强调发布数据集的重要性，这样，人工智能就可以用于解决社会难题了。潜在步骤可能包括：通过实施“人工智能公开数据”计划，实现公布大量政府数据集的目标，加速人工智能研究，在政府、学术机构和私营部门等领域，促进人工智能公开数据标准的使用和最佳实践。

**建议措施 3:** 联邦政府应该探索新方法，提高关键机构应用人工智能完成任务的能力。例如，联邦机构应该发挥潜能，设立类似国防高级研究计划局这样的机构，以支持高风险、高回报的人工智能研究及其应用。教育部就是一例，该部提出建立专项高级教育研究项目，以确定人工智能及其他技术是否能够显著提高学生的学习成绩。

**建议措施 4:** 美国国家科技委员会的机器学习和人工智能分会应该为政府的人工智能从业者们开辟一个实践社区。各机构应该共同努力，制定并分享人工智能在政府工作中的使用标准和最佳实践经验。各机构应该确保将人工智能引入到“联邦雇员培训计划”的可能性。

**建议措施 5:** 在制定人工智能驱动的产品政策时，相关机构应当吸收适当的高级技术专家。有效管理人工智能驱动的产品需要以下三方的合作：机构领导、了解现有管理框架和办法的雇员、人工智能技术专家。机构领导应当采取措施外聘必要的技术专家，或者从内部挑选专家，并且确保在管理政策讨论中保留足够的技术“讨论席位”。

**建议措施 6:** 机构应当充分使用人员任命和交流模式（如招聘部门），以培养联邦的工作人员，从多个角度了解当前的技术发展状况。

**建议措施 7:** 交通部应该与产业界和研究人员合作，加强用于安全、研究和其他目的的数据共享。未来，人工智能将在地面车辆和飞行器上得到重要应用。面对这样的未来，联邦机构应在短期内建立起丰富的数据集——收集数据时要注意保护消费者隐私——以便在这些技术成熟时能够做出更合理的决策。

**建议措施 8:** 美国政府应投资开发和应用一种高端的自动空中交通管理系统，该系统可扩展升级，也能够同时应对无人机和有人驾驶的飞行器。

**建议措施 9:** 美国交通部应继续致力于建立不断完善的制度框架，确保全自动车辆和无人飞行器的安全整合，包括将新型车辆设计融入到公共交通系统。

**建议措施 10:** 美国国家科技委员会机器学习和人工智能分委员会应该监控人工智能技术的发展，并定期向上级主管部门领导汇报人工智能的发展状况，尤其是关于其发展技术中的里程碑式突破。随着技术的不断发展和专家共识的不断变化，分委员会应及时更新里程碑清单，并考虑在合适的时机向公众报告人工智能技术的发展状况。

**建议措施 11:** 政府应该监控其他国家的人工智能发展状况，尤其是人工智能技术的里程碑式突破。

**建议措施 12:** 人工智能相关行业应该和政府合作，及时向政府反馈行业的最新发展动态，包括近期可能在技术上取得的重大突破。

**建议措施 13:** 联邦政府应该优先发展人工智能的基础和长期研究项目。如果联邦政府和私营企业能够在人工智能研发领域长期稳定地投入资金，尤其是在高风险的基础研究领域的长期投资，会让整个国家从中受益。由于私营企业尤其不愿意在长期的基础研究领域进行投资，这就更凸显出联邦政府在一领域进行研发投资的重要性。

**建议措施 14:** 美国国家科技委员会(NSTC)机器学习与人工智能网络信息技术研发分委员会，以及科学技术工程教育委员会应该联合发起一项关于人工智能从业者流水线的研究，以发展出更多事情及行动，促进从业者包括人工智能研究人员、专家和用户在数量、质量和多样性上的合理增长。

建议措施 15：总统办公室应在年底前出版一份后续报告，进一步研究人工智能和自动化对美国就业市场的影响，并提出应对政策的纲要。

建议措施 16：使用人工智能系统进行决策或为相关决策提供支持的联邦机构在决定个人事宜时，应当特别注意确保这些系统的效力和公平，并以证据为基础进行验证和确认。

建议措施 17：资助各州和地方政府使用人工智能系统进行有关个人的决策的联邦机构应当对资金的使用条件进行审核，确保使用联邦资金购买的以人工智能为基础的产品或服务能够产生足够透明的结果，并以证据确保其效力和公平性。

建议措施 18：学校和大学应该将道德、保障、隐私和安全等主题作为人工智能、机器学习、计算机科学和数据科学课程的一部分。

建议措施 19：人工智能专业人士、安全专业人士以及他们所在的专业团体应该共同努力将人工智能向着更为成熟的人工智能安全工程领域推进。

建议措施 20：美国政府应在涉及人工智能的国际参与方面制定政府层面的政策，并制作需要国际参与和监督的人工智能热点领域清单。

建议措施 21：美国政府应该加深与关键国际利益相关者的合作，包括外国政府、国际组织、业界、学术界等，以交流人工智能研发方面的信息并促进合作。

建议措施 22：机构的计划和战略应考虑到人工智能和网络安全之间的相互影响。参与人工智能问题的机构应该向美国政府和私营部门网络安全部门确保，人工智能系统和生态系统在应对智能对手时能够保持安全性和恢复力。参与网络安全问题的机构应该使美国政府和私营部门人工智能部门制定创新的方法应用人工智能，以实现有效和高效的网络安全。

建议措施 23：美国政府应该在符合国际人道主义法的基础上制定关于自动和半自动武器的统一政府政策。

#### 缩略语

|      |  |          |
|------|--|----------|
| AAAI | Association for the Advancement of Artificial Intelligence | 美国人工智能协会 |
| AGI  | Artificial General Intelligence                            | 强人工智能    |
| AI   | Artificial Intelligence                                    | 人工智能     |

|        |   |                         |
|--------|---|-------------------------|
| APEC   | Asia-Pacific Economic Cooperation   | 亚太经合组织                  |
| BRAIN  | Brain Research through Advancing Innovative Neurotechnologies                   | 使用先进革新型神经技术的人脑研究        |
| CALO   | Cognitive Agent that Learns and Organizes                                       | 具备学习和组织能力的认知主体          |
| CCC    | Computing Community Consortium  | 计算社区联盟                  |
| CCW    | Convention on Certain Conventional Weapons                                      | 《特定常规武器公约》              |
| CEA    | Council of Economic Advisers  | 经济顾问委员会                 |
| CEO    | Chief Executive Officer   | 首席执行官                   |
| CGC    | Cyber Grand Challenge (run by DARPA)  | 网络超级挑战赛(由 DARPA 举办)     |
| CoSTEM | Committee on Science Technology, Engineering, and Education (component of NSTC) | 科学技术工程教育委员会 (NSTC 下属组织) |
| CS     | Computer Science  | 计算机科学                   |
| DARPA  | Defense Advanced Research Projects Agency                                       | 美国国防高级研究计划局             |
| DoD    | Department of Defense   | 美国国防部                   |
| DOT    | Department of Transportation  | 美国交通部                   |
| FAA    | Federal Aviation Administration   | 美国联邦航空管理局               |

|       |   |                            |
|-------|---|----------------------------|
| FMVSS | Federal Motor Vehicle Safety Standards                              | 《美国联邦机动车安全标准》              |
| IARPA | Intelligence Advanced Research Projects Activity                    | 情报高级研究计划署                  |
| ICTs  | Information and Communication Technologies                          | 信息通讯技术                     |
| IPA   | Intergovernmental Personnel Act                                     | 《政府间人员交流法》                 |
| LAWS  | Lethal Autonomous Weapon Systems                                    | 致命性自主武器系统                  |
| MLAI  | Machine Learning and Artificial Intelligence (subcommittee of NSTC) | 机器学习与人工智能委员会 (NSTC 下属分委员会) |
| NAS   | National Airspace System  | 美国国家空域系统                   |
| NEC   | National Economic Council   | 美国国家经济委员会                  |
| NHTSA | National Highway Traffic Safety Administration                      | 美国国家公路交通安全管理局              |
| NIH   | National Institutes of Health                                       | 美国国立卫生研究院                  |
| NIPS  | Neural Information Processing Systems conference                    | 神经信息处理系统进展大会               |
| NITRD | Networking and Information Technology Research and Development      | 网络信息技术研发委员会 (NSTC 下属分委员会)  |

|      |   |                 |
|------|---|-----------------|
|      | (subcommittee<br>of NSTC)                                   |                 |
| NSF  | National Science<br>Foundation                              | 美国国家科学基金<br>会   |
| NSTC | National Science and<br>Technology Council                  | 美国国家科技委员<br>会   |
| OECD | Organization for<br>Economic Cooperation<br>and Development | 经济合作与发展组<br>织   |
| OMB  | Office of Management<br>and Budget                          | 美国行政管理和预<br>算局  |
| ONR  | Office of Naval<br>Research                                 | 美国海军研究总署        |
| OSTP | Office of Science and<br>Technology Policy                  | 白宫科技政策办公<br>室   |
| R&D  | Research and<br>Development                                 | 研发              |
| RFI  | Request For<br>Information                                  | 信息邀请书           |
| STEM | Science, Technology,<br>Engineering, and<br>Mathematics     | 科学、技术、工程与<br>数学 |
| UAS  | Unmanned Aerial<br>System                                   | 无人飞行器           |
| UTM  | UAS Traffic<br>Management                                   | 无人飞行器交通管<br>理   |