



美国卫生与公众服务部：人工智能在健康、公众服务和公共卫生中的应用战略计划

战略计划

2025 年 1 月

美国卫生与公众服务部



目录

致谢与免责声明	4
副部长致辞	5
引言	6
1. 医学研究与发现	18
1.1 引言与背景	18
1.2 参与医学研究与发现人工智能价值链的利益相关者	19
1.3 人工智能在医学研究与发现中的应用机会	22
1.4 医学研究与发现中人工智能的趋势	23
1.5 人工智能在医学研究与发现中的潜在应用场景及风险	24
1.6 行动计划	32
1.7 结论	49
2. 医疗产品开发、安全与有效性	51
2.1 引言与背景	51
2.2 参与医疗产品开发、安全与有效性的利益相关者	52
2.3 人工智能在医疗产品开发、安全和有效性方面的应用机会	54
2.4 人工智能在医疗产品开发、安全和有效性方面的趋势	55
2.5 医疗产品及其开发中人工智能的潜在用例和风险	56
2.6 行动计划	63
2.7 结论	77
3 医疗保健交付	77
3.1 引言和背景	77
3.2 参与医疗保健交付人工智能价值链的利益相关者	78
3.3 人工智能在医疗保健交付中的应用机会	82
3.4 人工智能在医疗保健交付中的趋势	82
3.5 医疗保健交付中人工智能的潜在用例和风险	83
3.6 行动计划	95
3.7 结论	105
4 公众服务提供	106
4.1 引言与背景	106

4.2 参与公共服务提供人工智能价值链的利益相关者	107
4.3 人工智能在公共服务提供中的应用机会	108
4.4 公共服务提供中人工智能的应用趋势	110
4.5 人工智能在公共服务提供中的潜在用例和风险	111
4.6 行动计划.....	116
4.7 结论	125
5 公共卫生	127
5.1 引言与背景	127
5.2 参与公共卫生人工智能价值链的利益相关者	128
5.3 人工智能在公共卫生中的应用机会	133
5.4 人工智能在公共卫生中的应用趋势	134
5.5 人工智能在公共卫生中的潜在用例和风险	135
5.6 行动计划.....	141
5.7 结论	153
6 网络安全和关键基础设施保护	155
6.1 引言和背景	155
6.2 参与健康和公共服务生态系统中网络安全和关键基础设施保护的利益相关者	156
6.3 网络安全和关键基础设施保护的的趋势	156
6.4 行动计划.....	158
6.5 结论	162
7 内部运营	163
7.1 引言和背景	163
7.2 机遇与风险	163
7.3 治理	164
7.4 内部流程改进与创新	166
7.5 劳动力与人才管理	167
7.6 结论	168
结论.....	170
附录 A: 术语词汇表.....	171
附录 B: 部分联邦政策和法规.....	189

致谢与免责声明

致谢

美国卫生与公众服务部（HHS）在此感谢 HHS 人工智能工作组、指导委员会、工作小组联合负责人以及撰稿人对本文件的贡献，他们付出了远超预期的大量时间和精力。我们也非常感激 HHS 众多同事提供的深思熟虑的意见，并参与了本战略计划的制定。此外，部门还要向 HHS 人工智能工作组项目管理办公室表达诚挚的谢意，感谢其领导、协调和指导工作。最后，HHS 感谢来自各界的众多利益相关者，他们自愿投入时间并分享观点，为本战略计划提供了信息。我们衷心感谢他们富有建设性和批判性的贡献。

免责声明

美国卫生与公众服务部的人工智能战略计划不会修改或解释《联邦食品、药品和化妆品法案》（FD&C 法案）、《公共卫生服务法案》、食品药品监督管理局（FDA）法规或其他相关法规下的任何要求。本文件也不构成 FD&C 法案第 701(h) 条 (21 USC. 371(h))、21 CFR 10.115 或其他条款所定义的指导文件。此外，本文件不会为任何公众成员确立任何权利或义务。

副部长致辞

人工智能（AI）对健康、公众服务和公共卫生产生了不可否认的影响。在美国卫生与公众服务部（HHS），我们一直坚定不移地致力于负责任地利用人工智能，以推进 HHS 内部关键领域以及整个行业的使命。

在 HHS，我们对人工智能的变革潜力持乐观态度。这些技术具备推动创新的非凡能力，可以加速科学突破、提高医疗产品安全性和有效性、通过医疗服务提供改善健康结果、增加公众服务普适性以及优化公共卫生。然而，我们的乐观态度伴随着强烈的责任感。我们需要确保美国人免受风险。人工智能的部署和应用应该使美国人民受益，我们必须让生态系统中的利益相关者承担责任以实现这一目标。人工智能为改善我国的健康和公众服务以及更好地服务美国人民创造了巨大机会，HHS 已经在积极采取措施，推动人工智能的道德和负责任使用，以改善人们的生活。

我们很高兴推出 HHS 人工智能战略计划，该计划展示了我们促进创新、推动可信人工智能发展、使技术和资源民主化以及培养人工智能赋能的劳动力和组织文化的举措。本计划是我们持续致力于利用人工智能的力量来增强我国健康和福祉的一个重要里程碑。

我们将按照本计划的详细内容，利用可用资源和手段，在健康、公众服务和公共卫生领域成功部署人工智能。但要取得成功，需要举国上下齐心协力，与行业、学术界、患者及众多其他群体展开合作。我们期待生态系统中的其他各方与我们一同肩负起这一使命。

副部长 安德里亚·帕尔姆



引言

计划目的

HHS 的愿景是在创新和使用负责任的人工智能方面成为全球引领者，从而在提升全体美国人的健康与福祉上取得前所未有的进展。HHS 人工智能战略计划（以下简称“战略计划”或“计划”）提供了一个框架和路线图，以确保 HHS 履行其对国家的义务，并率先负责任地使用人工智能来改善人们的生活。

在过去 50 年中，美国个人与数字技术互动的方式发生了深刻变化。人工智能对健康和公众服务既具有巨大的潜力，也存在潜在风险。虽然人工智能自 20 世纪中叶以来就以某种形式存在，但近年来已变得无处不在。这在医疗保健领域也是如此，并且在公众服务提供中也将越来越普遍。新兴技术使在症状出现之前预测疾病、识别有可能改变护理标准的新药物靶点以及更有效地将公众服务与最需要的人匹配变得更加可能。鉴于这项技术的发展轨迹，人工智能从根本上改变健康和公众服务的潜力将变得更大。

本战略计划将人工智能定义为《国家人工智能倡议法案》(15 U.S.C. 9401 (3)) 第 5002 (3) 条所概述的：一种基于机器的系统，能够针对给定的一组人类定义的目标，做出影响真实或虚拟环境的预测、建议或决策。在这个定义范围内，人工智能可以有多种形式。在医疗保健领域，用于执行任务或解决问题的基本算法已经广泛使用了几十年。人工智能和机器学习 (ML) 能力的进步正在加强算法，使其超越狭隘的规则，通过分析或“学习”可用数据变得更具预测性和通用性，从而更精确地根据特定个人或亚人群的特征定制模型输出。生成式人工智能 (GenAI) 是另一种人工智能，指的是通过分析和学习数据来创建（“生成”）新事物的技术，如数据、文本、图像、声音或其他类型的信息。

私营部门的兴趣和投资推动了人工智能和健康人工智能（用于健康和公众服务的技术）能力的快速增长。人工智能技术，尤其是生成式人工智能，在过去几年中发展迅速，仅在 2023 年，工业和学术环境就产生了 60 多个值得注意的机器学习模型。2024 年第二季度，风险投资和私人人工智能投资大幅增长，2024 年第二季度美国各行业的风险投资资金中，人工智能相关投资超过 550 亿美元。预计到 2032 年，生成式人工智能的投资将逐年增长高达 42%，从而在各行业形成一个潜在的 1.3 万亿美元的市场。在医疗保健领域，初创企业在过去三年中为人工智能筹集了约 300 亿美元。此外，为满足人口需求，公众服务提供方面也需要额外的投资（例如，世界卫生组织 (WHO) 估计，到 2050 年将有 35 亿人需要辅助技术，其中一些可能由人工智能驱动）。为确保人工智能的负责任使用，美国的相关实体提及人工智能的法规数量有所增加（2023 年为 25 项，比 2022 年增加了 56%）。全球、跨国和其他各国政府机构，包括联合国、七国集团 (G7)、经济合作与发展组织 (OECD) 和世卫组织，也同样将制定人工智能使用的指导方针和战略作为优先事项。

人工智能为越来越多的科学突破铺平了道路，在某些情况下，在图像分类和视觉推理等

任务中可能超越人类表现。人工智能还有可能极大地提高识别相关因素或预测结果的能力。此外，人工智能和机器学习的进步推动了在健康和公众服务“后台”中预测模型的更多使用，例如预约安排以及研究中的证据和文献综述。

人工智能已经或将会直接或间接影响每个美国人在健康和公众服务方面的体验。因此，人工智能的发展应使用“以人为本的设计”方法，确保其专注于为使用或接受人工智能支持服务的人们提供真正的益处。以下是一些将在下文更详细讨论的益处：

- 加速科学突破，可能提高生活质量和寿命
- 作为医疗产品的一部分或用于开发医疗产品，以提高安全性和有效性
- 通过医疗服务提供的创新改善临床结果并增强安全性
- 通过改进健康和公众服务福利提供，促进公平并赋予参与者权力
- 预测风险并迅速调动资源以预测和应对公共卫生威胁

这种潜力并非没有风险。虽然人工智能可以显著改善健康和公众服务的许多方面，但它也存在可能导致不良影响和结果的风险。根据数据和模型质量的不同，人工智能可能会产生不正确或不完整的输出。当重要决策部分或全部基于不准确的人工智能时，人们可能会受到伤害或被剥夺权利，资源也可能被滥用。此外，研究人员发现，人工智能可能会引入和传播偏差，这可能会错误分类人们的需求，对身心健康结果产生负面影响，并增加成本。负责任的人工智能使用还应确保公平获取和有益性，保护受保护信息，并在适用时获得适当同意，同时还要考虑对社会或环境的潜在意外负面影响。需要注意的是，这些风险和考虑因素可能因所使用的人工智能的复杂性而异（例如，简单的基于规则的算法与大型语言模型（LLMs）将有不同的考虑因素）。无论如何，健康和公众服务中的人工智能使用必须确保并对适当的人类监督负责，并且人工智能应被视为支持和辅助努力的工具，而不是现有问题的唯一解决方案。

联邦政府正在努力评估人工智能的潜力，同时确保其对所有美国人安全且公平。最大化机会和降低风险是 HHS 长期使命的核心：通过支持有效的健康和公众服务以及促进医学、公共卫生和社会服务基础科学的稳健、持续进步，提高所有美国人的健康和福祉。这一使命得到了我们的社区合作伙伴、州、部落、地方和 territorial（STLTs）政府、学术界和私营部门合作伙伴的支持和关联。这要求 HHS 继续协调努力和优先事项，以确保质量和安全，满足国家不断变化的健康和公众服务需求，同时找到鼓励创新和部署必要保障措施的平衡。同样，这需要赋予终端用户（包括患者、医疗服务提供者和其他人）权力，通过促进整个创新管道的合作，使他们能够塑造新技术如何负责任地融入他们的护理和服务。

人工智能能力、适用范围、易用性和应用速度的近期进步表明，它可能比预期更快、更深刻地影响健康和公众服务。HHS 及其运营和工作人员部门（“部门”）认识到在企业层面运营的价值和重要性，而不仅仅是在特定单位内为独立目的进行孤立使用。HHS 制定

明确的战略至关重要，以确保健康和公众服务组织能够根据一致的原则和目标充分利用人工智能。明确的战略对于管理人工智能投资组合以及确保 HHS 利用其部门之间的协同作用也必不可少。

HHS 在该领域发挥着至关重要的作用：它是研究与发现的投资者、健康产业的监管者、推动健康和公众服务创新的催化剂、医疗保健和公众服务的提供者以及患者安全、权利和隐私的保护者。鉴于在 HHS 职责范围内，各行业对人工智能的应用情况各不相同，因此需要采取一种负责任的方式来推动其开发与应用。HHS 将利用现有的监管结构来澄清指导方针，在必要时提供新的指导，并根据技术创新更新监督机制。在适当的情况下，HHS 还将寻求新的监管权力。尽管人工智能的不断发展可能会挑战监管模式，但 HHS 将继续利用所有可用的手段，包括政策、资金、教育和宣传等，以适应新的技术现实并支持健康和公众服务生态系统中的利益相关者。

计划的组织与使用

战略计划的组织

本战略规划专门致力于阐明 HHS 在健康、公众服务和公共卫生领域对人工智能的愿景和目标。作为美国政府最大的联邦机构之一，HHS 的部门设置和业务活动涵盖了卫生和公众服务的各个环节，从实验室研究到临床医疗服务，从药物研发到监测工作，从预防儿童贫困到为老年人及残障人士提供福利。鉴于其广泛的职责范围，本战略规划提出了一个统一框架，该框架由七个领域构成，旨在促进 HHS 涉及人工智能的各项政策、项目及活动保持协调一致。

- **主要领域** 代表了 HHS 价值链的特定部分，包括：
 - **医学研究与发现**：对疾病和生命过程的基本机制进行基础和临床前研究，将其转化为医疗创新和临床应用，并考虑其在整个医疗服务提供中的背景。
 - **医疗产品开发、安全与有效性**：药物、生物制品和医疗设备的开发、临床试验和监管批准、制造以及持续的安全和有效性监测。
 - **医疗服务提供**：为个人和人群提供医疗服务，以诊断、治疗、管理和预防疾病，并促进健康和福祉，以及提供支持这种服务的资金。
 - **公众服务提供**：为个人和家庭提供社会服务和援助，以满足他们在健康、福利、自给自足、安全和福祉方面的基本需求。
 - **公共卫生**：通过社会、组织、公共和私人社区以及个人的有组织努力和明智选择，预防疾病、延长寿命和促进人群的健康。
- **其他领域** 是跨越主要领域的功能区域，是实施战略规划所必需的：

- **网络安全和关键基础设施保护**：保护和推进对健康和公众服务功能至关重要的系统安全，以支持人工智能的使用。
- **内部运营**：HHS 部门用于内部运营和功能的政策、计划和基础设施，使 HHS 能够实施战略计划并适应快速的技术进步。

在每个主要领域内，各章节遵循一致的结构：

- 该领域人工智能的引言和背景
- 参与该领域人工智能价值链的利益相关者
- 人工智能的应用机会
- 人工智能的趋势
- 潜在用例和风险
- 行动计划

本计划的完整版本有意进行了扩展，以便为寻求更详细信息的读者提供背景和具体示例。它包括对机会、趋势、用例和风险的更全面讨论，以及详细的行动计划。如需高层次的概述，请参阅为提高对广大读者的可及性和实用性而编写的概述。

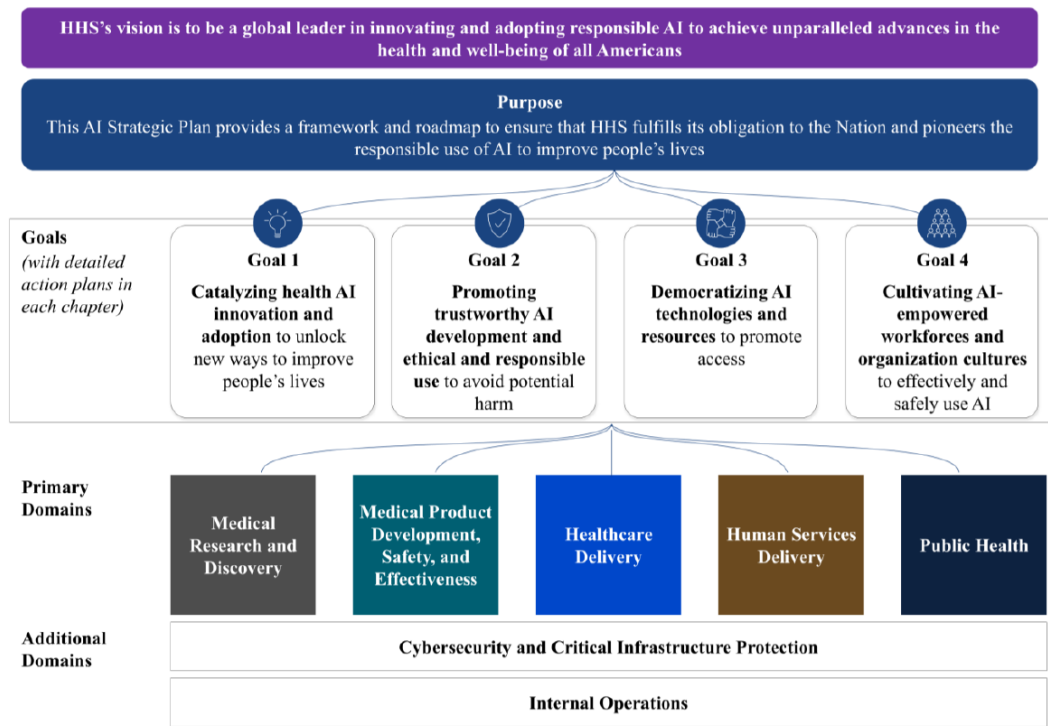
HHS 对战略计划的使用

HHS 的首要目标是通过创新、安全和负责任地开发和使用人工智能，启动一种协调的公私合作方法，以提高健康和公众服务的质量、安全、效率、可及性、公平性和成果。

HHS 将通过专注于以下四个关键目标来实现这一目标：

1. **促进健康人工智能创新与应用**，解锁改善人们生活的新途径；
2. **推动可靠的人工智能开发以及合乎伦理道德与负责任的使用**，避免潜在危害；
3. **普及人工智能技术和资源**，以提升获取；
4. **培养人工智能赋能的劳动力和组织文化**，有效且安全地使用人工智能。

Exhibit 1: Goals and Structure of the Strategic Plan



如附表 1 所详述，在每个主要领域内，各章节阐述了美国卫生与公众服务部（HHS）将如何通过当前及规划中的行动，聚焦这四个反复提及的目标，这些行动将指导并支持目标的执行。这些行动将涵盖 HHS 及其各部门可利用的多种手段，包括法规、政策与指南、拨款、资助项目、公众教育与宣传，以及内部基础设施建设、采购和运营等。需要注意的是，新政策并非支持负责任使用人工智能的唯一途径；现有方法也可进行更新，以应对新出现的问题，同时确保人工智能的使用仍符合现行法规（例如，患者隐私保护法规）。通过在其价值链中协调运用这些手段，HHS 旨在实现最大程度的协同，并从战略上使各部门以及卫生与公众服务生态系统的其他部分保持一致，以实现 HHS 的战略愿景，抓住人工智能改善人们生活的机遇。

人工智能为改善人们生活带来机遇

人工智能有潜力改善人们的生活，并在各个领域支持 HHS 更广泛的使命。以下是一些例子：

- **加速科学突破，最终提高生命质量与延长寿命：**自 2000 年以来，从一期临床试验到获得监管批准的平均时间约为十年，而基础研究和药物研发的前置时间甚至更长。在临床发现和开发的全过程融入人工智能，有望极大地助力聚焦安全有效的靶点，确定产品可能最为有效的人群和疾病，评估数据和数据模型的代表性，纠正人群抽样不足等问题，最终缩短开发周期并降低总体成本。
- **作为医疗产品的一部分或用于开发医疗产品，以提高安全性和有效性：**人工智能可作为医疗产品的一部分，或用于开发安全有效的医疗产品。特别是具备人工智能功能的

医疗设备，如非处方助听器，有可能被患者、医疗服务提供者及其他终端用户使用，以辅助提升医疗服务并改善治疗效果。此外，人工智能能够从临床使用过程中收集的数据中学习，这有助于随着时间推移提高医疗产品的准确性和性能，有可能带来更高的诊断准确性和更好的监测效果（例如，降低误诊率、提高早期发现不良反应的能力）。同样，人工智能可用于辅助开发药物和生物制品（例如，识别靶点、评估生物标志物和终点指标）。

- **通过医疗服务创新改善临床结果并增强安全性：**医疗失误，包括诊断错误和（或）延误，可能导致不良后果。人工智能有潜力通过快速处理广泛且分散的信息，发现人类观察不总能明显发现的模式，并通过临床决策支持及其他工具，针对个体情况引导临床医生做出更有可能正确的诊断和（或）关注安全问题，从而加快诊断速度并预防安全事件。人工智能还能优化医疗模式和健康服务研究，以开发出更有利于临床医生、支付方和患者的创新成果。
- **通过改善卫生与提供公众服务福利，增进公平性并赋予患者和公众权力：**如今，许多个人和社区因社会经济地位、语言、地理位置及其他因素，在获得医疗服务方面面临障碍。人工智能有潜力改善所有人获取福利和服务的机会；例如，对于那些因语言障碍而难以获得医疗保健或公众服务的个人，实时自动翻译技术提供的翻译服务可能会有所帮助。人工智能还能帮助残障人士完成简单或复杂的任务，比如语言技术可以通过优化语音模式，帮助有语言障碍的人进行流畅的对话。
- **预测风险并迅速调配资源，以预测和应对公共卫生威胁：**在应对诸如新冠疫情这样的公共卫生危机时，HHS 对人工智能的使用显著增加。大规模应用时，人工智能有潜力改善全球预测未来疾病爆发的基础设施，使公共卫生团队能够在新地区出现首例疾病之前，大规模制定有效的应对措施。人工智能还可通过其他方式改善公共卫生状况，例如识别可能影响卫生与公众服务的因素（例如，在自然灾害发生前进行预测，从而降低其影响）。

促进合乎伦理道德和负责任地使用人工智能

人工智能的使用也带来了一些固有的挑战和风险。HHS 致力于开发、分享和推广可信赖的人工智能，以改善健康和福祉结果。为支持这一承诺，HHS 正在确定确保人工智能可信的现有做法，并解决不同领域之间的不一致问题。虽然本计划的范围不包括针对每个潜在用例提出人工智能在健康领域道德和负责任使用的全面方法，但 HHS 在本计划中列出了适用于整个生态系统的总体考虑因素。HHS 期望所有组织最大限度地促进人工智能的道德和负责任使用。利益相关者应共同努力减轻无意伤害的风险，例如错误识别患者病情、违反患者信息的保密性（无论是直接还是通过重新识别加密和（或）去标识化的患者数据）、误导资源使用（特别是在公共卫生紧急情况下）、无意开发可能有害的医疗产品，或对社会或环境产生负面影响。利益相关者还应通过减少偏见和增加人群（如地理社区、残疾人）的获取机会来促进公平。

HHS 将基于现有的风险管理和治理框架，如美国国家标准与技术研究院 (NIST) 的人工智能风险管理框架和技术政策助理部长 / 国家卫生信息技术协调员办公室（以下简称“ASTP”或“ASTP/ONC”）的《健康数据、技术和互操作性：认证计划更新、算法透明度和信息共享 (HTI - 1) 最终规则》(89 FR 1192)。NIST 框架声称，全面的人工智能风险管理需要进行风险映射、测量和管理，以为行动和治理提供信息。HTI - 1 最终规则规定了一种风险映射方法，用于关键信息的透明度，以评估人工智能的益处和风险。NIST 和 HTI - 1 最终规则中确定的某些政策均基于 FAVES 原则（公平、适当、有效、安全）。

FAVES 原则如下：

- 公平：模型结果不会基于个人或群体的固有或获得特征表现出偏见或偏袒。
- 适当：模型和过程输出与应用它们的特定背景和人群相匹配，以产生适当的结果。
- 有效：模型和过程输出在内部和外部数据中均已证明能够准确估计目标值，符合预期。
- 安全：模型结果不存在任何已知的不可接受风险，且可能的益处超过可能的风险。

FAVES 并非涵盖所有可考虑的风险领域，但它的原则通过描述健康和公众服务背景下高质量人工智能的广泛特征，为评估人工智能的开发和使用提供了基础。本计划的各章节将讨论风险和缓解策略，以确保安全和可信赖的使用。随着人工智能的快速发展，HHS 将继续重新审视这些原则，并与利益相关者合作，以应对人工智能带来的挑战。所有人都有责任监测风险，并支持 FAVES 模型和人工智能的使用。

对州、部落、地方和地区卫生与公众服务组织的适用性

在许多情况下，人工智能被部署在各个州、部落、地方和地区 (STLT) 以及社区组织 (CBO) 中。HHS 认识到每个组织根据患者和人群健康因素都有独特的需求，并且在某些情况下，组织具有不同的责任（例如，一些 STLT 和 CBO 提供直接服务，而另一些则不提供）。HHS 将采取灵活的方法，支持创新，同时确保安全和负责任的开发与使用。通过这种方式，HHS 和行业合作伙伴可以在 STLT 和其他实体增加人工智能使用并确定改善健康和公众服务的新方法时向它们学习。相关实体和潜在行动在特定领域的章节中有更详细的讨论。

2024 年 4 月，HHS 发布了一份计划，旨在促进 STLT 政府在公共福利管理中对自动化和算法系统中人工智能的负责任使用。在该计划中，HHS 就 STLT 在公共福利和服务管理中应如何选择、采购、设计、治理和管理人工智能提供了建议。2024 年 4 月的计划还概述了 HHS 支持 STLT 制定自己在公共福利计划和服务的自动化和算法系统中使用人工智能的政策和实践的计划。HHS 在本战略中与这些建议保持一致，并描述了支持和促进 STLT 安全和负责任地开发和使用人工智能的其他优先事项。

HHS 与人工智能相关的角色和职责

为了与人工智能增强所有美国人健康和福祉的潜力保持一致，HHS 于 2021 年 3 月设立了首席人工智能官办公室，并设立了首席人工智能官 (CAIO) 一职。CAIO 位于 ASTP 内，其主要职责是推动战略计划的实施，监督 HHS 的人工智能治理结构，协调 HHS 对联邦人工智能指令的响应，并促进与人工智能相关的合作。CAIO 在 HHS 和联邦政府中发挥着至关重要的作用，以维持美国在人工智能领域的领导地位。在像 HHS 这样庞大而广泛的部门内履行对人工智能的承诺，需要各部门之间的协调以及全部门对负责任的人工智能原则和资源的统一。CAIO 将充当这一协调职能，整合不同部门的多样化能力，以推进战略计划。CAIO 还将监测部门间的交叉合作如何为健康和公众服务领域的人工智能创造新机会，填补更分散的策略可能遗漏的空白。ASTP 在 HHS 范围内的人工智能实施和应用的协调方面也将发挥更广泛的作用。

以下 HHS 部门将在评估人工智能机会方面发挥多种作用。以下是每个运营部门及其关键人工智能活动的简要描述：

- **儿童和家庭管理局 (ACF)**：管理 60 多个项目，提供福利和服务以支持家庭和儿童，包括促进经济和社会福祉。ACF 在 HHS 人工智能战略计划中的作用，将侧重于确保向儿童和家庭有效且公平地提供公众服务。
- **社区生活管理局 (ACL)**：支持为有复杂需求的人群（特别是老年人和残疾人）提供的项目，并管理包括营养服务、老年人支持服务和老年人权利等各种项目。ACL 在 HHS 人工智能战略计划中的作用，将侧重于确保向有复杂需求的个人有效且公平地提供公众服务。
- **医疗保健研究和质量局 (AHRQ)**：提供资金和项目，以提高医疗保健的质量、可及性、公平性、可负担性和安全性，包括改善初级保健和协助获得社会福利和公共卫生服务；管理和监督患者安全组织项目；授予研究人员发起的健康服务研究资金，包括数字医疗保健研究，如健康人工智能和临床决策支持；执行国家支出调查，收集利用、支出、支付来源和医疗保险覆盖范围等信息。在 HHS 人工智能战略计划中的作用，将聚焦于推动并开展相关研究，探索如何在工作流程中安全使用人工智能并合理应用，以实现高质量医疗服务。
- **卫生高级研究计划局 (ARPA - H)**：推进高潜力、高影响力的生物医学和健康研究，这些研究无法通过传统研究或商业活动轻易完成。ARPA - H 在 HHS 人工智能战略计划中的作用，将侧重于发布奖项，以催化前沿研究。
- **战略准备和响应管理局 (ASPR)**：领导国家应对灾害和其他公共卫生紧急情况的医疗和公共卫生准备、响应和恢复工作，并与医疗保健和公共卫生利益相关者（如 STLTs 和医院）及其他各方合作，提高国家的准备和响应能力。ASPR 在 HHS 人工智能战略计划中的作用，将侧重于在公共卫生紧急情况中协调人工智能的使用（与疾病控制和预防中心 (CDC) 及其他利益相关者合作）。
- **疾病控制和预防中心 (CDC)**：检测和应对新出现的健康威胁，开展研究，发布

指导意见，并设计解决国家最大健康问题项目，促进健康和安全行为、社区和环境，并培训公共卫生工作人员。CDC 在 HHS 人工智能战略计划中的作用，将侧重于研究人工智能在疾病预防中的功效，并在公共卫生工作中实施人工智能。

- **医疗保险和医疗补助服务中心 (CMS)**：管理医疗保险计划、医疗补助和儿童健康保险计划 (CHIP) 的联邦部分以及健康保险市场，这些项目共同为大约 50% 的美国人提供健康保险。此外，CMS 批准和监督项目豁免和示范项目，开发和测试医疗保健支付和服务提供模式，制定医疗保健服务提供者的健康和安全标准，实施质量倡议，并促进健康信息技术的使用和使用等其他职责。CMS 在 HHS 人工智能战略计划中的作用，将侧重于确定对人工智能支持的医疗服务的适当覆盖范围（利用支付和监管政策确保支付者和提供者对人工智能的可靠、负责任使用），监督和认证州信息技术系统和数据收集标准，并向提供者、州和其他利益相关者提供技术援助。
- **食品和药物管理局 (FDA)**：通过在上市前后评估其安全性和有效性来监管医疗产品（包括药物、生物制品和医疗设备）。FDA 还通过促进创新来推进公共卫生，这些创新有助于加速患者获得安全、有效和创新的医疗产品。FDA 还肩负着维护我国食品供应（供人类和动物食用）、化妆品以及辐射类产品安全的职责。此外，FDA 还对烟草制品的生产、营销及分销进行监管，以保护公众健康。在 HHS 的人工智能战略中，FDA 的作用将集中于制定基于风险的方法，对具备人工智能功能的医疗产品以及用于研发医疗产品的人工智能进行监管；发布行业指导意见；加强与国际监管机构的监管合作。
- **卫生资源和服务管理局 (HRSA)**：为国家需求最高的社区提供公平的医疗保健，包括通过支持低收入人群、艾滋病病毒感染者、孕妇、儿童、父母、农村社区、移植患者和医疗工作者的项目。这包括在 HRSA 支持的健康中心接受护理的 3100 多万人、5800 多万孕妇、婴儿和儿童、56 万多艾滋病病毒感染者、全国 1900 多个农村县和市以及通过贷款偿还和奖学金项目支持的近 22000 名医疗服务提供者。HRSA 在 HHS 人工智能战略计划中的作用，将侧重于确保公平使用人工智能以使服务不足的社区受益，并教育和培训未来几代医疗专业人员。
- **印第安卫生服务局 (IHS)**：为部落民族和社区提供初级和急性护理，通过在印第安保留地或附近的 600 多家医院、诊所和卫生站网络，为大约 280 万美国印第安人和阿拉斯加原住民提供服务。IHS 在 HHS 人工智能战略计划中的作用，将侧重于在这些人群中实施人工智能医疗服务，并确保人工智能指南对相关 STLTs 的适用性。
- **国立卫生研究院 (NIH)**：开展和资助生物医学研究，并为旨在改善国家健康的项目提供领导和指导。NIH 在 HHS 人工智能战略计划中的作用，将侧重于开展和资助研究以推动生物医学、行为和健康研究中的人工智能发展，制定和评估必

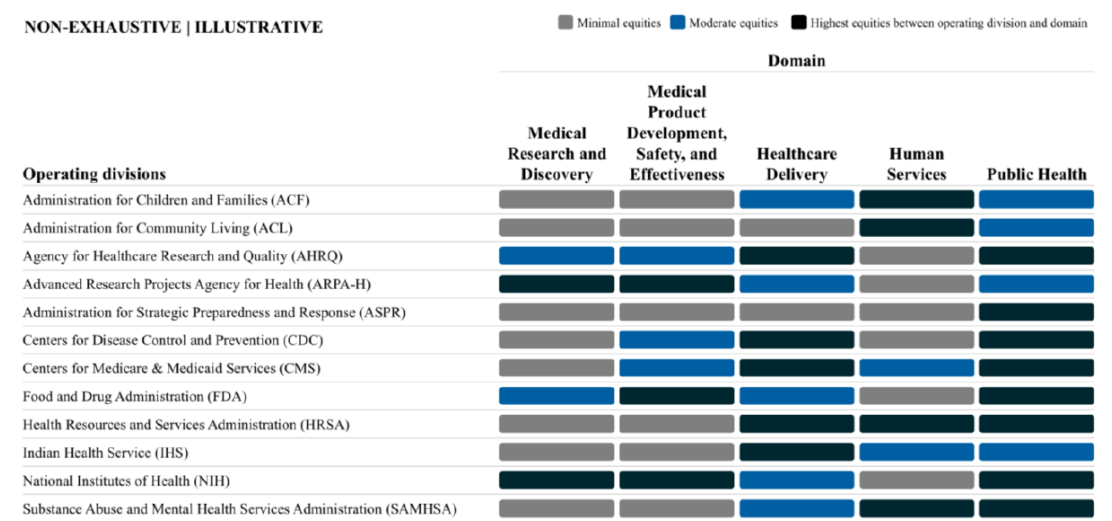
要的标准，支持开发人工智能模型培训的最佳实践，培养和培训人工智能人才，并促进人工智能的负责任使用。

- **物质滥用和心理健康服务管理局（SAMHSA）：**领导努力减少精神和物质使用障碍对个人、家庭和社区的影响。SAMHSA 提供资金、指导和资源以支持预防、治疗和康复服务，确保公平获得护理。SAMHSA 在 HHS 人工智能战略计划中的作用，将侧重于向 STLT 社区提供赠款资金和指导，并收集、分析和分发行为健康数据以评估项目、改进政策并提高对预防、减少危害、治疗和康复资源的认识。SAMHSA 还将支持行为健康临床医生和卫生系统使用人工智能。

HHS 各部门在许多领域具有互补和相互依存的责任。虽然运营部门可能跨越多个领域，但以下示意图大致概述了每个领域中各部门的权益：

附录 2：卫生与公众服务部运营部门的权益概览

注意：此示意图仅指示哪些部门参与哪些领域，需要进行协调与合作。它并非旨在详尽说明每个部门的权益，并且部门可能以不同方式在不同领域发挥作用。



除了上述运营部门外，HHS 的工作人员部门将在确保战略计划成功方面发挥重要作用。例如，ASTP 负责监督医疗保健、公共卫生和公众服务中临床信息的访问、交换和使用的数据和技术标准的实施。它还指导各种联邦项目中对健康信息技术（如电子健康记录）的监管，并支持政府和行业参与者之间实现互操作性。ASTP 的职责将集中于 HHS 跨部门政策制定，以及协调人工智能的实施与应用。民权办公室（OCR）执行联邦民权法（如第 1557 号最终规则）、良心和宗教自由法、《健康保险流通与责任法案》（HIPAA）隐私、安全和违规通知规则以及《患者安全法案》和规则，这些法律共同保护非歧视、良心、宗教自由和健康信息隐私等基本权利。OCR 的作用将是在整个人工智能开发和使用过程中提供保护个人权利的教育。规划和评估助理部长办公室、全球事务办公室和卫生助理部长办公室也有其他相关职责。首席信息官办公室在支持 HHS 内部人工智能的使用方面也将发挥显著作用。这并非 HHS 所有内设部门及其全部工作的详尽清单，

而是一种用以突出该部门广泛的工作流程和职责，并阐明协调工作重要性的方式。HHS 各部门的单独行动以及协作努力，对于落实本战略规划至关重要。

行动计划摘要

以下章节将阐述支持这些目标的现有和计划中的活动。这些行动被组织成主题，详细说明了 HHS 对人工智能未来的期望，如下表所示：

行动支持的关键目标	各章节行动主题（非详尽，详细行动计划见各章节）
1. 促进健康人工智能创新与应用，解锁改善人们生活的新途径	- 扩大每个领域价值链中人工智能的使用范围 - 现代化基础设施以实施人工智能并支持使用 - 加强合作和公私伙伴关系以促进人工智能使用 - 澄清人工智能的监管监督和覆盖 / 支付决定过程 - 支持收集人工智能干预措施（如疗效、安全）的结果证据和最佳实践
2. 推动可靠的人工智能开发以及合乎伦理道德与负责任的使用，避免潜在危害	- 建立和传播支持减轻公平、生物安全、数据安全和隐私风险的证据 - 制定明确标准，指导在可信人工智能使用背景下联邦资源的使用 - 支持人工智能风险管理的组织治理 - 完善监管框架以应对自适应人工智能技术 - 促进外部评估、监测和透明度报告，并促进其他健康人工智能质量保证机制
3. 普及人工智能技术和资源，以提升获取	- 增加对合理管理的数据和基础设施的访问，包括在适当情况下为组织提供支持 - 支持信息共享机制，传播标准、最佳实践并促进合作以改善访问 - 开发用户友好、可定制和开源的人工智能工具 - 增强州、部落、地方和地区（STLT）及其他社区组织的能力，包括在适当情况下提供资源或其他机制
4. 培养人工智能赋能的劳动力和组织文化，有效且安全地使用人工智能	- 改善人工智能治理和管理方面的培训 - 培养和保留强大的人工智能人才队伍 - 为专业人员提供获取资源和研究的途径，以支持其各自的健康和公众服务组织 - 利用人工智能缓解劳动力短缺并解决职业倦怠和人员流失问题

HHS 的愿景是在创新和使用负责任的人工智能方面成为全球引领者，从而在提升全体美国人的健康与福祉上取得前所未有的进展。本战略计划的以下章节详细说明了实现这一愿景的具体行动。

1. 医学研究与发现

1.1 引言与背景

医学研究与发现通过推动创新药物、生物制品、医疗设备（包括一些基于软件的行为干预措施）以及其他改善个人和社区健康结果及获得优质护理的工具的发展，对促进健康至关重要。

本章将重点关注医疗产品的研究与发现以及可在生物医学中应用的人工智能技术的研究与发现。医疗产品生命周期的后续阶段，包括临床试验，以及其他领域的研究，如卫生系统、公众服务提供和公共卫生，将在其他章节中讨论，不在本章范围内。

近年来，医疗技术和制药公司、学术和研究机构以及其他组织越来越多地利用人工智能来支持其医学研究与发现活动，并创建人工智能驱动的工具，但目前尚未充分利用现有的人工智能技术的全部潜力。虽然进一步的发展可能会带来更多好处，但需要采取行动来促进人工智能的安全和负责任使用，以更充分地实现其在医学研究与发现环境中的潜力。因此，本章解释了行业趋势、人工智能用例和风险，以及 HHS 可以采取的有助于在医学研究与发现中，安全启动人工智能的行动。HHS 提供了医学研究与发现的高层次背景以及该领域人工智能的概述，包括所涉及的利益相关者和人工智能使用的关键机会。

医学研究与发现为评估用于人类疾病诊断、预防、缓解及治疗的诊断方法、疗法、治疗手段、疫苗、技术及其他工具提供了数据和信心。从宏观层面看，它们可在一条价值链中加以描述，该价值链包含三个阶段：基础研究、发现（不同类型的医疗产品在这一阶段情况可能有所不同）以及临床前研究。关于这条价值链及其各组成阶段的详细讨论，见下文 1.5 节“人工智能在医学研究与发现中的潜在用例及风险”。

在医学研究与发现的各个方面，美国卫生与公众服务部（HHS）在激发相关活动、促进安全性与质量方面发挥着积极作用。美国国立卫生研究院（NIH）近 83% 的资金用于资助院外研究及研究支持；此外，NIH 遵循 HHS 通用规则，并制定了自身政策，以确保人体研究对象的安全，维护数据安全与质量，并为参与研究的弱势群体提供额外保护。其他部门也发挥着变革性作用：2023 财年，美国高级研究计划局卫生与公众服务部（ARPA - H）和医疗保健研究与质量局（AHRQ）的预算分别为 15 亿美元和 3.74 亿美元，以推动生物医学和健康领域的突破性创新。此外，美国食品药品监督管理局（FDA）对旨在生成证据以支持研究性药物（人用和兽用）、生物制品及医疗器械安全性和有效性的科学研究进行监管。尽管这只是概括了 HHS 部分部门在医学研究与发现中的作用，但还有许多其他部门以不同方式参与其中。随着人工智能在医学研究与发现中应用日益广泛，HHS 及其核心参与部门将推动在整个生态系统中安全且有效地使用公平的人工智能技术。

1.1.1 行动计划概述

本章稍后部分，美国卫生与公众服务部（HHS）将阐明为推进该部门在负责任使用人工智能方面的四个目标而拟采取的行动。以下是每个目标下行动主题的概述。有关拟议行动的完整详细内容，请参阅 1.6 节“行动计划”。

行动所支持的关键目标	拟议行动的主题（并非详尽无遗，更多细节见 1.6 节行动计划）
1. 推动健康领域人工智能创新与应用	- 扩大人工智能在医学研究与发现中的应用广度，涵盖不同疾病领域及价值链各环节。 - 加强跨地区协作，利用人工智能改善医学研究与发现。 - 培育适用于人工智能的数据标准和数据集，增强其在基于人工智能的医学研究与发现中的可用性。
2. 促进可靠的人工智能开发以及合乎伦理与负责任的使用	- 生成并传播相关证据，降低生物安全、数据安全、隐私及数据收集风险。 - 为医学研究与发现中安全可靠地使用人工智能，以及联邦资源的分配与使用制定明确指南。 - 实现对人工智能风险管理和透明度的安全且负责任的组织治理。
3. 实现人工智能技术与资源的普及化	- 促进有针对性的公众参与及公私合作行动，加强所有利益相关者之间最佳实践的分享。 - 提高所有人对经过负责任整理的适用于人工智能的数据、模型与算法，以及工具和基础设施的可及性。
4. 培育具备人工智能能力的劳动力队伍与组织文化	- 改进医学研究与发现中人工智能治理与管理方面的培训。 - 培养并留住医学研究与发现领域强大的人工智能人才储备。

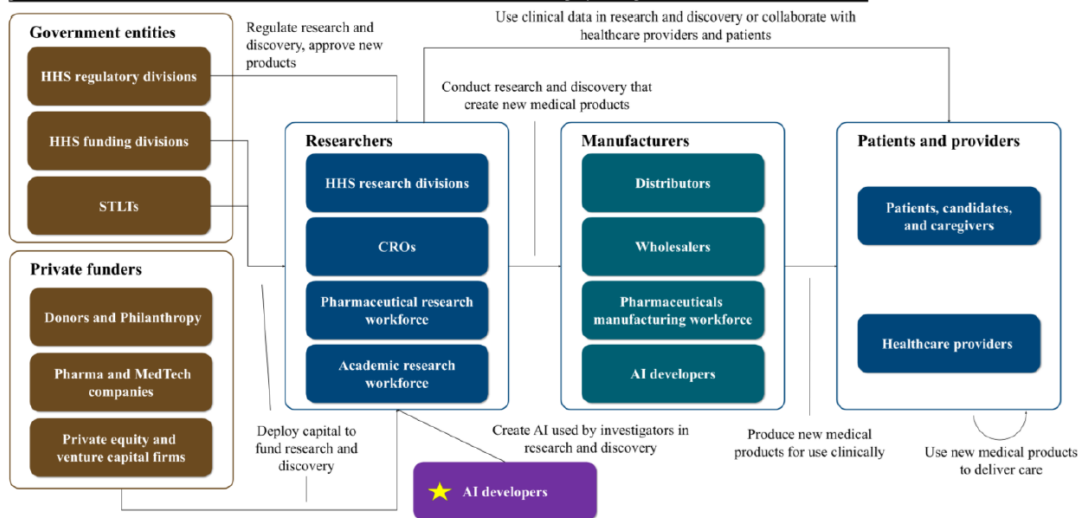
1.2 参与医学研究与发现人工智能价值链的利益相关者

医学研究与发现最终必须满足当前及未来患者及其护理人员的需求；因此，相应的人工智能应用应推动能满足这些需求的研究及最终技术的发展。除了患者和医疗服务提供者，还有几个关键利益相关者在医学研究与发现中涉及人工智能的运用，范围从医疗产品开发者到分销商、服务提供者、支付方、研究人员等等。本章末尾的“行动计划”部分包含了与这些利益相关者合作的方法，以便在推进创新的同时降低风险。以下是一个说明利益相关者之间示例流程的示意图，以及一份带项目符号的列表，其中包含有关医学研究与发现利益相关者的更多详细信息。请注意，该示意图和列表均未涵盖所有可能的利益相关者角色及互动。有关更多监管指导和权力细节，请参考美国卫生与公众服务部（HHS）的其他文件。

Exhibit 3: Stakeholders Engaged in Medical Research and Discovery

NON-EXHAUSTIVE | ILLUSTRATIVE

Below is an example of the flow of authorities and actions in medical research and discovery. It does not capture all the permutations and intricacies of stakeholder roles and interactions, and entities can play multiple roles.



★ Note that there are also AI developers in the academic and pharmaceutical research workforce

Please see information on official FDA, OCR, ASTP/ONC, NIH, and other HHS websites for more detailed information on regulatory authorities in medical research and discovery.

利益相关者（包括合作伙伴）包括：

- **美国卫生与公众服务部（HHS）下属业务部门（列举不全）：**参与医学研究与发现领域人工智能相关工作的部门包括：
 - 美国国立卫生研究院（NIH）：支持美国国内及国外的生物医学与行为研究，在其自有实验室和诊所开展研究，培养有前途的年轻研究人员，并推动生物医学知识的收集与共享。近年来，这些活动越来越多地涉及与医学研究和发现相关的人工智能技术（例如，提供数据，在生物医学研究与发现中催生数据科学与人工智能应用机会，增加人工智能模型开发的多样性，以及在生物医学研究各领域开发和应用人工智能）。
 - 美国高级研究计划局卫生与公众服务部（ARPA - H）：通过支持开发针对社会上最具挑战性的健康问题的高影响力解决方案，来加快改善所有人的健康状况，其中包括利用人工智能的解决方案（例如，使用人工智能加速抗生素的发现与开发）。
 - 美国食品药品监督管理局（FDA）：帮助确保人用和兽用药品、生物制品以及医疗器械在预期使用中的安全性和有效性，同时确保发射辐射的电子产品的安全性。随着人工智能在医学研究与发现中变得越发重要，FDA 将继续在监管产品和支持利益相关者方面发挥作用。
 - 医疗保健研究与质量局（AHRQ）：致力于通过研究、技术评估以及推广和实施工作，提高所有美国人医疗保健的质量、安全性、效率和有效性。AHRQ 将专注于促进和开展关于安全使用人工智能以实现高质量医疗服务的研究，传播可

行的、基于证据的人工智能知识，并提供覆盖决策所需的证据。

- **其他联邦机构：**HHS 还与许多其他联邦部门密切合作，如国家科学基金会（NSF）和能源部（DOE）。
- **患者、研究参与者、护理人员和相关倡导团体（包括居民和社区）：**历史上，他们被视为各种类型人工智能设计或嵌入的诊断、治疗、疗法、疫苗、技术和其他工具的接受者或管理者。虽然以患者为中心并非新事物，但现在有能力的患者也可以利用人工智能更好地了解自己的个人健康状况并倡导自己的护理；他们可以参与研究和发现过程（例如，作为研究早期规划阶段的合作者）。
- **学术、非营利和其他研究人员：**研究人员致力于推动生物医学知识前沿的发展，工程师设计和制造用于临床应用的医疗设备，以及主题专家开发人工智能、在研究工作流程中应用人工智能和（或）将人工智能集成到产品开发周期中。他们是医学研究与发现中人工智能的主要用户之一。
- **制药、生物技术和医疗设备行业研究人员：**负责设计、开发和生产用于医疗保健提供的诊断、治疗、疗法、疫苗、技术和其他工具，包括将人工智能集成到研究工作流程和产品设计中的研究人员和主题专家。他们也是医学研究与发现中人工智能的主要用户之一。
- **医疗服务提供者：**医院、诊所和医疗专业人员利用医疗产品，经常参与医学研究与发现，以提供临床视角。此外，提供者可以作为医学研究与发现价值链中的“人工环节”。
- **州、部落、地方和地区政府（STLTs）：**联邦政府以外的监管机构。虽然医疗产品由 FDA 监管，但医疗实践通常由 STLTs 管辖。此外，STLTs 可以资助医学研究与发现活动。
- **分销商和批发商：**促进医疗产品（可能是通过利用人工智能进行研究和发现的产品）向医疗服务提供者的分销。
- **合同研究组织（CROs）：**提供外包研究服务，可能更集中于临床开发，这在医疗产品开发、安全与有效性章节中有详细阐述，并且可能在其医学研究与发现价值链或工作流程中开发或集成人工智能。
- **捐赠者和私人资助者：**非营利捐赠者（如基金会）和营利资助者（如私募股权、风险投资和其他资助组织）通过支持上游研究资金在医学研究与发现及持续发展中的作用。这些组织也可能支持对数据集聚合、平台或人工智能工具开发以及在过程中使用人工智能的直接投资。
- **人工智能优先技术开发者：**构建可在整个医学研究与发现价值链中使用的人工智能工具（如蛋白质折叠软件）、模型、数据基础设施和平台（如电子健康记录）的工程师和组织。开发者包括人工智能优先的生物技术公司、大型科技公司和特定领域的参与者。

HHS 将让利益相关者参与与医学研究与发现中人工智能相关的任何资金机制、政策指南、教育材料或内部基础设施的开发或完善，以确保 HHS 促进这些技术的获取、理解和潜在影响的公平性。此外，与 STLTs 密切合作，特别是其卫生和公众服务监管机构，

将使本计划在各级政府和各地域保持一致。让生态系统中的利益相关者参与将对执行这项工作至关重要。

1.3 人工智能在医学研究与发现中的应用机会

在医学研究与发现价值链中负责任地使用和扩展人工智能有可能通过以下方式改善美国人的健康结果和获取途径：

1. **增强基础研究获得新的生物学见解，以改善人类健康的潜力：**并非所有医学研究与发现都是直接“转化性的”（即旨在产生可立即在医疗护理中采取行动的结果）。事实上，“基础”研究（即旨在更深入地理解现象或机制）在历史上对人类健康产生了一些最具影响力的下游影响（例如 CRISPR）。通过利用人工智能结合临床使用数据（例如在长寿研究中）检查疾病与核心病理过程之间的联系，基于对大量数据的快速分析探索更多假设，筛选图像以辅助人类研究，并高速生成见解，新的基础研究发现不仅可能增加，而且可能比没有人工智能支持时质量更高。最重要的是，这种转变可能会带来更好的人类健康。
2. **增加可及性以推动创新，并可能降低成本：**新出现的证据表明，在医学研究与发现价值链中利用人工智能带来了每年高达 260 亿美元（仅药物方面）的财务机会，对于医疗器械可能还有额外价值。如果得以实现，这种效率可能会降低开展医学研究与发现的障碍，并（或）释放资金用于进一步的医学研究与发现活动。例如，医学研究与发现成本可能在很大程度上受“湿实验室”房地产的驱动，在该空间中可以测试物理生物和化学样本，其每平方英尺的租金可能几乎是传统办公空间的两倍。通过利用人工智能在计算机上进行医学研究与发现的某些步骤（如蛋白质折叠建模、生物相互作用模拟），对湿实验室空间的需求可能会减少，这可能会降低创新所需的成本。人工智能可以使资金获取渠道较少的机构（如初创企业、非营利组织、学术研究机构）参与创新，增加医学研究与发现的多样性，从而可能带来更多突破。此外，如果将这些潜在的成本降低进行再投资，可能会激发更多机会。虽然不同产品的成本和时间线各不相同，但例如一些药物的总开发成本可能在 3 亿美元到 45 亿美元之间。如果每年 260 亿美元的潜在财务机会得以实现并再投资，这可能会大大加速新创新的可用性。
3. **扩大医学研究与发现的范围，以满足未被满足的患者需求并支持突破性创新：**人工智能可能促进突破性创新以及针对历史上服务不足患者的健康需求的新型医疗产品的开发。鉴于单个医疗产品的研究与发现相关的巨大成本和时间，今天的研究与发现活动可能侧重于潜在更有利可图的治疗领域（TA），而不是最有健康需求的领域。通过利用人工智能将研究与发现扩展到此类“稳妥”目标或疾病之外，并在其他 TA 中追求突破性创新的同时增加对潜在研究不足的 TA 的研发活动，可能会改变服务不足患者的健康结果和获取途径。通过平衡“值得探索”的靶点或治疗领域范围，人工智能还有可能减少基础医学研究与发现中的偏差。

4. **加速新产品研发和医疗服务可及性：**目前，药物的临床前开发尤其估计需要六到十年时间。然而，近年来，在医学研究与发现中利用人工智能在相应用例中显示出了希望（例如，从人类需要数年确定蛋白质结构到仅需几秒钟）。如果人工智能在医学研究与发现价值链中成功且负责地被使用和扩展，这些效率可能会显著缩短医疗产品到达患者手中的时间，拯救美国人的生命，改善健康结果，并更快地惠及服务不足的患者。作为医学研究与发现的世界领导者，美国也可能在全球范围内加速医疗产品的获取。

1.4 医学研究与发现中人工智能的趋势

医学研究与发现中，人工智能的使用正在增长，遵循以下几个关键趋势：

1. **人工智能在医学研究与发现价值链中的使用在增加但不一致：**截至目前，人工智能在研发中的应用更多集中于确定性活动，特别是在靶点识别和先导化合物生成方面（例如，预测蛋白质折叠、分子相互作用以及细胞疾病过程）。具体而言，在计算机模拟环境下对生物分子进行设计、操控和探索，以及对设备进行设计等方面，可能比基础研究或临床前研究的应用场景，更多地使用了人工智能技术（有关价值链中各应用场景使用人工智能的示例，详见下文“人工智能在医学研究与发现中的潜在应用场景及风险”部分）。
2. **人工智能的应用可能集中在具有更强市场激励的治疗领域（TAs）：**研究人员面临强大的激励因素，比如诱人的知识产权所有权，这使得他们将医学研究与发现活动聚焦于有利可图的治疗领域。而人工智能的应用未必能解决这一情况，甚至可能加剧这种趋势（例如，更多的数据会带来更好的模型，这些模型又被用于对盈利性治疗领域的进一步研究与发现）。人工智能投资可能也面临类似的激励，倾向于关注与探索“高置信度靶点”相关的应用场景，这可能包括专注于已知而非新的靶点。然而，通过正确的干预措施来克服这些结构性激励因素，人工智能可被用于研究较少的靶点和治疗领域，并实现满足未被满足的患者需求的突破性创新。正如上文 1.3 节“机会 3：扩大医学研究与发现的范围，以满足未被满足的患者需求并支持突破性创新”所强调的，这是一个巨大的机会。
3. **医学研究与发现正在超越传统实验室：**虽然研究人员利用人工智能加速医学研究与发现，但其他参与者（如技术公司）也带着新颖的人工智能创新进入研究与发现生态系统。例如，确定蛋白质的动态结构曾经需要晶体学，这是一个通过 X 射线衍射使蛋白质结晶以阐明其原子位置的艰苦过程，且需要使用湿实验室空间。最近的投资催生了一种开源算法，能够预测许多蛋白质的结构以及它们如何在体内折叠和与其他蛋白质及分子相互作用。现在，一些实验在计算机上可以显著更快地完成。然而，这些初步结果仍应通过生物方法进行验证，或者需要有人参与其中以确保准确性。在技术公司追求此类解决方案的同时，制药和医疗技术公司也在构建人工智能应用

程序，这些应用程序可用于改变医学研究与发现价值链中任务的质量，并加快完成这些任务的时间。

4. **数据碎片化且基础设施成本上升：**人工智能在医学研究与发现中的成功应用需要大量高质量的训练数据，这对机器学习和其他模型的基础至关重要。如今，大约 75% 包含可用于医学研究与发现人工智能模型数据的学术文献都在付费墙之后（随着公共访问政策的实施，这一情况可能会改变）。对于医学研究与发现可能非常有用但不存在于“学术记录”中的大量潜在数据是碎片化的，难以汇总和整理（例如真实世界数据）。此外，利用人工智能所需的专用硬件和计算成本高昂，且能耗高。资源较少无法获取此类技术的实体可能会因价格过高而被排除在外，这阻碍了公平使用并限制了创新。如果在医学研究与发现中不能公平、安全地访问数据，这些限制将更加严重。
5. **自主人工智能和其他自主系统可能会不断发展：**HHS 致力于以符合道德和安全的方式使用人工智能，包括可能自主人工智能。虽然目前处于萌芽阶段，但具有自主问题解决和协作能力的自主人工智能系统有望成为实验室的一部分，以帮助增强研究人员在整个医学研究与发现价值链中的活动。与遵循编程规则的传统人工智能不同，自主人工智能可以独立或协作地进行分析、决策和行动。自主人工智能可以使医学研究与发现更快，并使突破性创新更快地惠及患者。HHS 已经在采取行动以应对这一趋势；例如，美国卫生高级研究计划局（ARPA - H）已经发布了一份信息请求，以了解自主人工智能并为医学研究与发现设定其相应的战略方向。

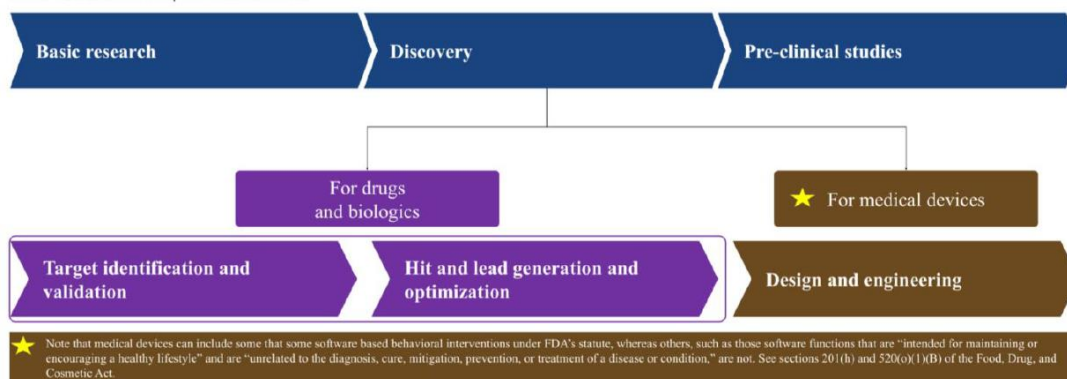
1.5 人工智能在医学研究与发现中的潜在应用场景及风险

医学研究与发现价值链

在美国，医学研究与发现是一个严格的多步骤过程，旨在增强生物学知识，并确保药物、生物制品和医疗设备在进入市场之前的安全性和有效性。虽然可能存在差异，但一般来说，它形成了一个三步价值链：（1）基础研究，（2）发现（不同类型产品的步骤有所不同），以及（3）临床前测试。临床试验（如有相关内容）将在“医疗产品开发、安全与有效性”章节中讨论，不在本章范围内。同样，关于卫生系统、护理模式和非医疗设备行为干预的研究不在本章范围内，而是包含在“医疗服务提供”章节中。此外，医学研究与发现活动的这个价值链可以在迭代反馈循环中为其他领域（如公共卫生、医疗服务提供和公众服务提供）提供信息。

Exhibit 4: Medical Research and Discovery Value Chain

NON-EXHAUSTIVE | ILLUSTRATIVE



1. **基础研究**涉及科学探索，可揭示生物学、疾病或行为的基本机制，以增进对生物现象和可观察事实的一般知识或理解，这对人类健康的进步至关重要，也是美国国立卫生研究院（NIH）资助基础研究的原因之一。一个领域前沿的小步前进可能会为开发人员带来新的生物标志物或作用机制作为目标，并使研究人员和公众有信心最终在价值链的这一步骤之外与人类研究参与者一起测试新药物、生物制品、医疗设备、技术和其他工具。
2. **发现**是诊断、治疗或治愈疾病的科学探索，根据医疗产品类型不同而有所不同，如下所述：
 - a. **对于药物和生物制品**（如治疗剂、疫苗）：
 - i. **目标识别与验证**对药物开发的早期阶段很重要，这通常依赖于初步识别适合药物候选物的生物目标。这包括找到治疗剂可以调节的生物系统（如神经回路、内分泌或免疫途径），并确保该目标的参与具有“潜在治疗益处”。如果一个目标无法验证，它将不会进入药物开发过程。
 - ii. 命中和先导物生成与优化识别具有所需生物活性的化合物或其他治疗类型，这些化合物或治疗类型与已验证的目标结合可产生预期的治疗反应。随后进行优化，在改进结构缺陷的同时保持先导化合物的有利特性。这一步骤的目标是确定一种用于临床前测试的化合物。
 - b. **对于医疗设备**（如诊断设备、一些行为干预措施，如下所述）：
 - i. **设计与工程**是创建新设备概念或想法的过程。在此基础上，研究人员确定判断该概念是否可行所需的步骤。然后通过原型构建，进一步完善该概念。注意：一些基于软件的行为干预措施根据 FDA 法规属于医疗设备，而其他一些，如那些“旨在维持或鼓励健康生活方式”且“与疾病的诊断、治疗、缓解、预防无关”的软件功能则不属于。见《联邦食品、药品和化妆品法案》（FD&C 法案）第 201(h) 和 520(o)(1)(B) 条。有关非设备行为干预措施的研究信息，请参阅“医疗服

务提供” 章节。

3. **临床前测试**是指体外和体内研究，旨在进一步推进用于人类临床研究的潜在治疗方法。通常这样做是为了在人体试验之前确定任何毒性或不良影响，并最终使其能够在市场上销售。如果一种药物或设备显示出潜在益处，研究人员可以向 FDA 提交研究性新药申请（药物）或研究性设备豁免申请（设备）以进入临床试验，这在“医疗产品开发、安全与有效性” 章节中有更详细的讨论。

医学研究与发现中的人工智能风险

由于医学研究与发现是产品使用和护理提供的前序步骤，在这些步骤中使用的人工智能模型所产生的任何偏差或其他未考虑到的风险都可能向下游传播，潜在地影响到患者。因此，考虑、管理并最终减轻相关的人工智能风险至关重要。此外，如果在患者、护理人员 and 提供者眼中，人工智能在研究和技术方面没有建立起可信度，可能很难实现大规模应用。随着技术的快速发展，积极主动地让这些群体参与进来对于促进这些技术的安全使用可能是必不可少的。虽然潜力巨大，但未来的成功将取决于关键参与者如何共同努力平衡风险和管理不确定性。

在本章后面详细阐述医学研究与发现中人工智能的额外益处和风险之前，先强调三个管理风险的重点领域：生物安全、数据安全和人工智能劫持。需要注意的是，这些风险尚未完全理解，并且可能随着技术的进步而演变，这使得难以对它们与其他风险进行分层和优先级排序。

1. **生物安全风险**：2024 年 5 月，总统执行办公室发布了《美国政府对两用研究关注和具有大流行潜力病原体的监督政策》，该政策阐明了人工智能两用应用的潜在情况。这包括为合法目的进行的研究，这些研究产生的知识、信息、技术和产品可用于改善护理结果，以及为恶意目的进行的研究，这些研究可能产生潜在有害的生物武器或有害病原体，对美国和世界构成生物安全威胁。帮助减轻这种威胁的行动已经开展（见行动计划部分的详细信息），未来，HHS 和美国政府安全机构可以继续与学术界、私营公司（包括制造商）和出版业密切协调，在现有总统执行办公室指导的基础上，继续努力在开放科学和公共安全之间取得适当平衡。
2. **数据安全风险**：2024 年 10 月的白宫《推进美国在人工智能领域领导地位的备忘录》指出了医学研究与发现中的一些特定风险：在这个过程中使用的人工智能系统可能会泄露其训练数据的某些方面——无论是无意的还是恶意行为者故意操纵的结果——当在不允许此类信息的网络上使用时，可能会导致模型数据泄露，这些模型可能是基于机密或受控信息进行训练的。未来，HHS 将探索需要哪些政策和技术支持，以确保在人工智能研究和开发中负责任和安全地使用这些数据。
3. **人工智能劫持**：恶意行为者可以在医学研究与发现的背景下劫持人工智能模型和系统，通过控制代理或解决方案将其导向有害行动。这可能与基础研究中分析大型生物医学数据集或药物和设备的设计与操纵等人工智能用例特别相关。人工智能劫持

可能包括污染训练数据。由于人工智能劫持可能导致个人健康信息、受控或机密信息以及专有或国家安全信息的泄露，因此它是一种跨领域风险，因此在下面的表格中不会在每个用例中列出。

1.5.1 医学研究与发现价值链中人工智能的使用示例和风险

在下表中，HHS 突出显示了医学研究与发现价值链中人工智能的非详尽潜在益处和风险列表。需要注意的是，下面详细介绍的用例强调了该领域各种利益相关者目前或潜在使用人工智能的方式。有关 HHS 及其部门如何使用人工智能的详细信息，请参考《2024 年 HHS 人工智能用例清单》。

功能组件 1：基础研究

推进对生物现象和可观察事实的一般知识或理解

潜在用例（非详尽）	潜在风险（非详尽）
先进的生成和分析模型，可加速突破性发现的时间线并扩展潜在假设的清单 <i>例如，分析医学文本和其他数据源以生成新的生物学见解</i> 利用人工智能对来自现有科学研究、出版物及其他数据源的大量信息进行分析与整合 <i>例如，对大型生物数据集库进行分析，以创建并完善可供探究的假设</i> 对大型数据集进行深度处理，以便更好地理解某种病症、生物机制或其他健康	偏差和有效性 - 可能引入偏差或产生不准确结果 例如，由于分析有偏差或低质量数据而得出不可推广的见解 人工智能驱动的基础研究成果，其质量可能仅与所分析的数据相当。如果数据集不能充分代表总体人群，研究结果可能无法推广应用。这种偏差随后可能会在医学研究与发现价值链的其余环节传播，甚至会影响到临床试验及其他方面所使用的医疗产品。此外，还可能通过数据投毒对数据或模型质量进行恶意操纵的情况，攻击者会篡改训练数据，致使人工智能“以不良方式运行”，这可能会影响结果的有效性和准确性。
	隐私、安全与透明度 —— 存在潜在的机密、敏感、保密或个人数据泄露，以及未经授权的披露风险。 <i>例如，有意或无意地泄露个人或机密信息，或对其进行重新识别</i> 人工智能模型有可能使用机密或其他敏感数据进行训练，这可能会带来原本应保密的信息泄露风险。举个例子，如果训练数据包含受保护的健康信息（PHI），如临床图像和（或）医疗记录，数据泄露可能导致用于

议题，这能够提高取得突破性发现的可能性 <i>例如，分析可能与疾病相关的基因、RNA（核糖核酸）和蛋白质</i> 基础模型能够分析海量基因数据，并借助机器学习来识别哪些生物分子可能与疾病有关	训练的 PHI 被提供给人工智能用户，从而引发潜在的监管和政策问题（例如，违反《健康保险流通与责任法案》(HIPAA)）。此外，随着模型收集和分析的数据量增加，即使数据最初是经过去识别化处理的，恶意行为者（有意）甚至算法（无意）重新识别出知情或不知情参与者的风险也会增加。在整合多个数据集或模型时，原本在各个数据集中已去识别化的数据，合并后可能会再次变得可识别。再者，当一个人工智能模型在一次分析中使用了已获得患者授权的 PHI，却在后续未经患者授权的人工智能分析中意外或故意使用该数据时，就会出现同意问题。未来，此类风险可能需要新的同意和授权框架，以及更高的透明度。 <i>例如，对于可能包含个人数据的临床数据如何用于基础研究缺乏透明度</i> 在基础研究中使用或为基础研究提供参考的人工智能模型，可能会使用可识别或已去识别化的患者数据（例如，用于训练疾病模型）。其数据可能被使用的人，可能并不清楚他们的数据是如何被使用或披露的，这种使用和披露可能产生的相应影响，以及任何伴随的风险。随着人工智能在基础研究中的应用不断增加，关于数据使用的适当授权和透明度机制将变得越来越重要。 <i>例如，随着数据集和模型的整合，模型卡片出现不准确的情况</i> 实现人工智能透明度的一种方法是利用模型卡片，它用于描述模型质量（例如，关于人口统计学、时间、数量和地域的训练数据情况）。随着模型和（或）其相关数据集的整合，相应的模型卡片可能会失去准确性，因为相互关联的数据和模型会增加与隐私、信息重新识别等相关的风险。请注意，这种风险可能适用于价值链的多个环节，但由于基础研究中人工智能用例涉及大量数据集，所以在此处进行阐述。
--	--

功能组件 2：发现

科学探索以寻找可能治疗或治愈疾病的疗法或开发产品，因医疗产品类型而异。

有关医疗产品类型更多详细信息，请参阅上述“发现”部分的描述。

潜在用例（非详尽）	潜在风险（非详尽）
预测模型，可利用基础研究见解预测和优先考虑潜在治疗靶点和先导物 例如，分析系统生物学以预测靶点 使用先进的分析方法对结构和系统生物学知识以及来自健康人和患有特定疾病的人的可用基因组、转录组、蛋白质组和其他数据源进行分析，以预测新靶点 例如，分析药物 - 靶点相互作用，通过药物重新定位来助力发现新疗法。 对药物 - 靶点相互作用进行探索，有助于预测可能与相同靶点相互作用或具有相似作用机制的各类药物，这或许能基于特定已知特征预测分子的毒性。此策略可指导药物重新定位工作，即利用先前已表征的化合物。利用人工智能开展药物重新定位工作，还有可能受益于从各种来源（如电子健康记录（EHRs）、登记处和数字健康技术（DHTs））获取的更丰富的合适真实世界数据，以此识别药物对疾病通路的未知影响。 例如，推荐研究靶点和先导化合物。 计算机辅助药物设计使研究人员能够快速预测抗体结构、评估氨基酸诱变的结构和功能，并加速从头蛋白质设计（例如，通过生成式人工智能验证肿瘤学靶点）。 例如，设计具有特定期望功能的核酸和氨基酸序列。	偏差和有效性 - 可能引入偏差或产生不准确结果 例如，基于非代表性数据集的靶点识别与先导化合物生成，以及潜在的人工智能社会偏见。 使用质量欠佳或不具代表性的数据集进行训练的模型（例如，从种族或性别比例失衡的人群中获取的生物标志物和生物分子数据），可能导致识别出仅适用于部分人群的靶点和先导化合物，这有可能延续社会偏见，加剧健康不平等，对特定群体造成伤害。尽管模型已学会如何改进因训练数据产生的偏差，但研究表明，潜在的偏差即便没有更严重，至少也同样普遍存在，这会加剧健康不平等，且更难追踪。 例如，源于异质或错误数据的统计和计算偏差。 在人工智能系统中，用于开发人工智能应用的数据集和算法流程可能存在统计和计算偏差。当算法基于某一种数据类型进行训练，且无法对该数据类型之外的情况进行推断时，就会产生这种偏差。误差可能源于数据的异质性、复杂数据在简化数学表示中的呈现、错误数据、诸如过拟合和欠拟合之类的算法偏差、异常值的处理，以及数据清理和插补因素。 例如，化合物或设备的识别不准确。 一些人工智能（如大语言模型）生成的内容，从设计原理上可能直接基于某些

利用人工智能平台创建具有有益功能的生物分子，能够提高药物研发的效率和速度。	信息，或间接推断得出（常被称为“幻觉”），这就可能导致将不准确的内容呈现为准确的，有时甚至会生成更多不准确的信息，在进一步追问解释时为这些不准确之处自圆其说。如果使用的人工智能在算法上没有致力于减少这种现象（例如，通过检索增强生成模型），可能会给医学研究和发现流程带来这种风险，并且如果不能以其他方式妥善解决（例如，引入人工干预），不准确的内容会在整个价值链中传播。 <i>例如，资源不必要地消耗在毫无根据的靶点或先导化合物上。</i> 人工智能在靶点识别、命中及先导化合物生成与优化相关的分析和预测中出现的幻觉或其他不准确情况，可能会将资金和（或）计算资源消耗在那些可能不适合进一步探索的靶点或先导化合物上。
计算机模拟实验技术能够预测行为、设计与操控产品，以及筛选有效候选药物。 <i>例如，蛋白质折叠预测助力产品设计。</i> 借助深度学习，基于大量数据存储库来预测蛋白质结构的模型。 <i>例如，生物分子与医疗器械的设计和操控。</i> 在临床前研究之前，对生物分子（如 DNA、RNA 和蛋白质）的结构进行计算机模拟实验，以测试候选药物和作用机制；或者对医疗器械的结构进行模拟实验，以帮助确定其潜在适用性。	生物安全威胁 —— 存在制造有害产品的潜在风险 <i>例如，恶意或无意设计出新型致病或有毒的生物及化学制剂，包括核酸序列、蛋白质和多肽。</i> 在公开可得的研究数据上运用人工智能，或者利用人工智能驱动的设计与折叠技术，既可能出于正当目的，也可能被用于恶意用途。与不使用人工智能的传统研究活动相比，利用人工智能可能更容易产生新型致病或有毒制剂，而现行的研究监管框架，如 2024 年美国《两用性关注研究及具有增强大流行潜力病原体监督政策》，尚未涵盖这些制剂。这些制剂的 DNA 和 RNA 序

例如，药物化合物筛选。 基于化合物对靶点的特异性和亲和力，预测化合物的化学性质、生物活性、功效及潜在不良事件。	列，也可能无法通过美国科学技术政策办公室（OSTP）《核酸合成筛查框架》中的现行最佳匹配标准检测出来。此外，其他制剂（如蛋白质、多肽）可能会突破自然免疫系统或现有的疾病治疗医学干预手段。
---	--

功能组件 3：临床前测试

在能够对人体开展试验之前，通过在细胞和（或）动物模型中对药物、医疗程序或医疗器械进行研究，以确定其是否存在任何毒性或不良影响。

潜在用例（非详尽）	潜在风险（非详尽）
预测模型、分析设备及展示工具，这些能够加快获得治疗的时间线，并在进入试验阶段前加深对研究成果的理解。 例如，预测药物和器械的疗效与安全性，以确定是否适合开展试验。 在临床试验前，基于多模态数据（如登记数据、组学数据、知识图谱、真实世界数据）对潜在疗效进行比较，从而降低风险，并有可能为潜在的失败节省大量试验成本。 例如，对体内和体外测试进行医学影像分析。 对研究图像进行自动分析，以识别结构，帮助选择用于临床试验的药物。 例如，利用数字孪生技术增加体内和体外测试的多样性及样本量。 物体、系统或动物候选对象的虚拟表示，通过开展更多模拟测试，能够加	有效性、偏差与效能，包括潜在的假阳性和假阴性问题 例如，无意中传播无效的想法或摒弃有前景的解决方案 如果没有人工介入来评估对数以百万计（甚至更多）已识别的潜在药物、器械和研究对象的分析有效性，那么在结果综合与优先级排序过程中出现的错误，可能会导致推荐结果出现假阳性和假阴性。 例如，随着合成数据的不断迭代，模型完整性和多样代表性下降 使用合成数据，即便初衷是为了增加多样性，但在对其进行反复分析以生成更多合成数据的过程中，可能会削弱模型质量。这可能危及结果的准确性和有效性，最终无法实现增加代表性和（或）减少偏差的潜在目标。

快并强化临床前研究。 <i>例如，生命科学工作流程优化。</i> 机器学习可用于“预测数百万种工作流程配置”，并对其进行优化，使其在分布式计算数据基础设施上尽可能高效地运行，从而加快研究发现。	研究人员技能退化 <i>例如，人工主导的实验室流程减少</i> 可靠、安全的实验室程序和操作实现自动化生成，可能会降低处理高风险生物材料所需的技能和培训要求，这可能导致重要的高技能人才流失。
大语言模型 (LLMs) 能够提升监管申报流程的质量与速度。 <i>例如，生成式与分析式的监管文件撰写。</i> 利用生成式人工智能 (GenAI)，依据临床前研究成果来开发用于研究性新药申请以及其他需要大量文字撰写工作的临床前试验步骤的申请材料。 <i>例如，借助数字助手实现程序与分析自动化。</i> 在实验过程中，智能代理助手对科学记录的输出进行维护、分析和综合（例如，环境监听、“人工智能科学家”）。	研究结果可能缺乏可解释性 <i>例如，监管申报材料未能准确体现临床前研究的成果。</i> 在人工智能技术中，并非总能追溯到模型所使用的原始数据，这会降低其输出结果或中间结论的可验证性。这种潜在的有效性缺失，可能会降低利益相关者（如学术界、产业界、公众和监管机构）对结果的信任。

有机会开发和应用人工智能来提高医学研究与发现的质量、数量和速度。从支持在实验室通过靶点识别和优化聚焦假设的人工智能，到分析大型数据集，有充分的理由保持乐观。然而，这种热情必须与现实相平衡，即这些应用存在风险，值得仔细关注和采取缓解策略。每个利益相关者都必须监测和缓解风险。HHS 将使用以下行动计划来授权价值链中的实体和个人更安全、负责任、公平和有效地使用人工智能。

1.6 行动计划

鉴于医学研究与发现中不断变化的人工智能格局，HHS 已经采取了多项措施来促进负

责任的人工智能使用，包括为内部和外部研究提供资源、提高精简数据集的可访问性、开发劳动力人才和能力等。以下行动计划遵循支持 HHS 人工智能战略的四个目标：1. 促进健康人工智能创新与使用；2. 推动可信人工智能开发和道德与负责任的使用；3. 使人工智能技术和资源民主化；4. 培养人工智能赋能的劳动力和组织文化。对于每个目标，行动计划提供了背景、HHS 和其他相关联邦行动的概述，以及 HHS 将采取的近期和长期具体优先事项。HHS 认识到随着技术的发展，该行动计划需要不断修订，并致力于提供结构和灵活性，以确保长期影响。

1.6.1 促进健康人工智能创新与应用

增加医学研究与发现中人工智能的应用可以改变创新的质量和速度，最终改善患者结果。HHS 有机会通过以下行动主题来增加人工智能的使用：

1. 扩大医学研究与发现中人工智能在疾病领域和价值链步骤中的使用范围
2. 加强跨地域协调以利用人工智能改进医学研究与发现
3. 培养人工智能就绪的数据标准和数据集以增强其对人工智能赋能的医学研究与发现的可用性

以下，HHS 更详细地讨论了每个行动主题的背景、迄今为止的相应行动以及促进医学研究与发现中人工智能创新和使用的计划。

1. 扩大人工智能在医学研究与发现中，横跨各类疾病领域以及价值链各环节的应用广度：

背景：

与价值链的其他环节相比，人工智能在研发环节（如计算机辅助靶点识别、潜在候选药物的高通量筛选）的应用程度相对较高，再加上人工智能受潜在激励因素影响，倾向于聚焦市场潜力较大的疾病领域。这表明，若能关注价值链其他环节（如基础研究、临床前研究）的人工智能应用场景，以及探索更多疾病领域（如研究较少、未满足需求高的领域），将有机会推动人工智能得到更广泛应用。HHS 专注于扩大人工智能在医学研究与发现中的应用，同时保持其资源计划的完整性，这可能包括资源分配、培训或额外的政策或指南。HHS 将着眼于推动人工智能的应用，以帮助满足未满足的患者需求，并在整个价值链中更广泛地促进创新。

HHS 迄今为止的行动（非详尽）：

- **美国国家癌症研究所 (NCI) 的癌症研究信息技术 (Informatics Technology for Cancer Research)** 资助跨越整个开发周期的研究驱动信息技术，以解决癌症研究中的优先需求。这些项目越来越多地开发或整合先进的人工智能方法。该计划支持开发关键工具和资源，以改善癌症研究连续体中数据的获取、分析、可视化和解释，包括癌症生物学、癌症治疗和诊断、早期癌症检测、风险评估和预防、癌症控制和流行病学以及癌症健康公平性。

- **美国国家心理健康研究所（NIMH）的理论与计算神经科学计划（Theoretical and Computational Neuroscience Program）**支持专注于基于生物物理现实的计算方法的基础实验和理论研究，这些方法用于模拟大脑中的动态过程，从单个细胞活动到调节复杂行为的神经系统。
- **NIMH 的转化数字和计算精神病学计划（Translational Digital and Computational Psychiatry Program）**促进创新的计算方法，以识别和验证预防和治疗精神疾病的新机制、生物标志物和治疗靶点。该计划支持使用先进计算方法结合行为、生物和（或）临床数据的研究项目，以破译精神疾病中涉及的复杂机制，并对预测风险、临床轨迹和治疗反应的新工具进行初步测试。
- **ARPA - H 的 TARGET 项目**将使用深度学习筛选候选生物分子并利用生成式人工智能拓宽可能的药物范围，从而扩大具有抗生素潜力的候选分子库。
- **ARPA - H 的计算 ADME - Tox 和生理学分析促进更安全治疗项目（CATALYST）**设想未来可以基于计算机模拟的安全数据批准开始首次人体临床试验。该项目专注于开发无动物、可靠的实验实践方法，特别关注药代动力学（包括吸收、分布、代谢和排泄（ADME））和药效学的安全性和毒性。CATALYST 将追求能够可靠代表人类生理学的新技术，以降低研究性新药候选物的失败率。这些技术将确保进入临床试验的药物具有可靠的安全性概况，并更好地保护不同的试验参与者和未来患者。
- **NIH 的通过推进创新神经技术进行大脑研究计划（BRAIN）**：从大脑复杂数据进行分析的理论、模型与方法项目，旨在开发相关理论、计算模型以及分析工具，以便从复杂的神经科学数据中深入理解大脑功能。拟议项目可以开发工具来整合现有理论或制定新理论；构建概念框架来组织或融合数据以推断大脑功能的一般原则；创建多尺度/多物理模型来生成新的可测试假设以设计/驱动未来实验；开发新的分析方法来证实关于大脑功能的可证伪假设。开发的工具预计将在神经科学研究社区中广泛可用并可修改。

HHS 近期优先事项：

- 探索为利用人工智能解决有未满足需求的治疗领域和（或）识别和分析新的而非已知的靶点的医学研究与发现提供资源。
- 探索为专注于基础和临床前研究领域（人工智能使用率较低，如接近监管批准过程的后期研究）的研究、培训和研讨会提供资源。
- 继续举办网络研讨会、研讨会、听证会等，以宣传资助机会通知（NOFOs）和信息请求。
- 识别人工智能在价值链中使用的障碍。
 - 召集利益相关者界定在医学研究与发现价值链中使用人工智能的技术、

经济、劳动力、数据可用性和监管障碍。

- 召集患者和其他利益相关者解决透明度问题并建立信任（见“促进可信人工智能开发和道德与负责任的使用”下的“实现人工智能风险管理和透明度的安全和负责任组织治理”）。
- 探索降低使用障碍的潜在机制（例如，与使用相关的环境考虑因素和成本）。
- 优先考虑并探索为证据建设提供资源，以评估负责任的人工智能医学研究与发现投资并最大化 HHS 支出的效果。
- 为联邦资助的临床前研究中人工智能的可接受用途提供政策清晰度和（或）指南（例如，使用人工智能替代基于动物的研究）。
- 为使用人工智能起草研究资助申请和提交提供政策清晰度和（或）指南，以确保公平性和透明度并保护项目完整性。
- 在 HHS 内使用人工智能以简化资助审查、批准和支持流程，同时采取强有力的保障措施保护项目完整性、公平性和公正性。

HHS 长期优先事项：

- 探索关于交换数据和人工智能模型的经济框架的实验机会，以使定价可负担，同时允许公平补偿和人工智能使用的安全性。

2. 加强跨地域协调以利用人工智能改进医学研究与发现：

背景：

国际和美国的多个机构有不同的法规，可能会影响医学研究与发现领域（例如，通用数据保护条例、欧盟人工智能法案和 HIPAA）。这些机构在医学研究与发现背景下对人工智能的方法进行协调，可以降低创新障碍，同时保持相应用例的安全性和有效性。如果没有积极的协调，考虑到价值链的复杂性、潜在经济因素以及大量公共和私营部门利益相关者的参与，在这些领域实现可实现的改进将是分散和不理想的。HHS 可以通过与国内外利益相关者合作，促进价值链的进一步协调，从而推动未来的创新。

HHS 迄今为止的行动（非详尽）：

- NIH 共同基金的利用数据科学促进非洲健康发现与创新（Harnessing Data Science for Health Discovery and Innovation in Africa, DS - I Africa）项目利用数据科学技术和先前的 NIH 投资，通过来自学术、政府和私营部门的新合作伙伴组成的强大生态系统，开发解决非洲大陆最紧迫公共卫生问题的解决方案。

HHS 近期优先事项：

- 优先考虑并探索为基础和临床前研究中最有前途的协作、多学科和跨境人工智

能集成提案提供资源。

- 促进 HHS 部门之间的协调, 以共享与医学研究与发现相关的适当数据、方法、技术和资源, 从而实现更强有力的 HHS 创新活动。

HHS 长期优先事项:

- 定义并建立符合美国标准的医学研究与发现跨境人工智能协作政策和指南。
- 向其他机构和 STLTs 提供与人工智能和数据共享标准相关的指南, 在 HHS 领域内适当且授权的情况下, 以增强医学研究与发现领域国际协作的可能性。

3. 培养人工智能就绪的数据标准和数据集, 以增强其对人工智能赋能的医学研究与发现的可用性:

(欲了解更多关于数据基础设施和工具的信息, 见目标 3: “普及人工智能技术与资源”, 行动 2 的主题: “让目前获取渠道较少的群体, 更易获得经过妥善整理、适用于人工智能的数据工具和基础设施”)

背景:

用于训练人工智能的数据在质量、数量和代表性方面的差异, 这可能因偏差和“走捷径”式学习导致人工智能表现不佳。虽然医疗服务体系会生成海量的临床和管理数据, 但该行业的碎片化状况给高质量数据的整合带来了极大挑战, 而这些数据对于开发支持临床前医学研究与发现的人工智能模型至关重要。此外, 基于包含个人信息的临床数据训练出来的模型, 很难广泛共享。而且, 专有或机密的分子、化学和其他非临床数据可能分散在行业和学术界。因此, 大量本可用于研究的数据难以轻易取得。HHS 通过专注于使这些数据适用于医学研究与发现领域的人工智能应用, 从而推动该领域对人工智能的进一步应用。

HHS 迄今为止的行动 (非详尽):

- **NIH 的 Bridge2AI 项目**为多项研究提供资金, 以生成具有示范意义的数据集, 并建立用于收集和准备适用于人工智能的数据的最佳实践方法, 从而应对生物医学与行为研究领域的挑战。(例如, 生成新的、符合伦理来源的、可靠的、定义清晰且易于获取的生物医学与行为示范数据集; 开发软件和标准, 统一来自多个数据源及多种数据类型的数据属性; 创建自动化工具, 加快生成符合 FAIR 原则 (可发现、可获取、可互操作、可复用) 且来源符合伦理的数据集; 提供资源, 以传播数据、伦理原则、工具及最佳实践方法; 制作培训材料并开展相关活动, 推动人工智能、生物医学和行为研究领域之间的人才培养。)
- **NIH 开发了 ScHARE**, 这是一个基于云的数据平台, 包含联合的社会决定因素健康 (SDOH) 数据集, 以加速健康差异、医疗保健提供、健康结果和人工智能偏见缓解策略方面的研究。

- NIH 与美国国家科学基金会（NSF）、国家科学与工程统计中心（NCSES）以及美国能源部（DOE）共同参与了**国家人工智能研究资源（NAIRR）试点项目**。在此项目中，NIH 借助大规模真实世界数据（RWD）集来开展以下工作：（1）构建一个合成数据生成器工具包和框架，以评估使用此类数据进行证据构建时的隐私风险及效用；（2）将医学影像数据与临床记录相链接，为多模态人工智能的发展增强能力。
- **NIH 的大脑研究计划（BRAIN）：数据档案库**通过创建一个具备适当标准和概要信息的数据档案库来推动研究进展。该档案库广泛面向研究界开放，便于其开展进一步研究。团队与研究界合作，整合了软件工具，使用户能够分析和可视化数据，并使用适当标准对数据进行描述。
- **NIH 的大脑研究计划（BRAIN）：整合与分析该计划的数据**，开发了信息学工具，用于分析、可视化和整合与 BRAIN 计划相关的数据，以增进我们对大脑的理解。这些工具便于用户访问和分析来自相应数据档案库的数据，用户无需下载数据，即可进行数据的分析与可视化操作。

HHS 近期优先事项：

- 与美国政府合作伙伴（例如，由于其 2024 研究数据框架（RDaF）而与 NIST 合作，以及与美国核心互操作性数据（USCDI）合作）定义并优先确定最大化研究数据（包括公共数据元素、元数据、持久标识符和安全）的可查找性、可访问性、可互操作和可重用性的标准，以简化使用生物医学研究数据进行算法训练和优化的过程。
- 加速联邦资助的研究数据标准（语义、格式、传输）与 HHS 使用的电子健康记录（EHRs）、医疗保健提供者和支付者标准（例如，USCDI、USCDI +、HL7 快速医疗保健互操作性资源（FHIR）、CARIN）的对齐。
- 开发用于人工智能数据管理、跨标准数据映射、去标识化等的开源、开放标准工具和基础设施，以开发人工智能就绪数据集和工具。
- 加速与标准开发组织的合作以及行业在标准方面的协作，以支持人工智能在整个生命周期中的开发和使用。
- 召集一个公私合作实践社区，分享关于适合医学研究与发现中人工智能模型使用的数据的最佳实践，利益相关者还可以在合作中确定访问此类数据的促进因素/障碍。
- 探索潜在的安全方法，在不危及隐私、征得同意以及保证透明度的前提下，利用并分享基于临床或其他个人信息训练的人工智能模型。
- 加速联邦机器学习研究、工具和实施支持；促进公私合作过程以定义联邦机器学习的开放行业标准和惯例。

- 加速在可信交换框架和通用协议（TEFCA）中开发研究交换目的，以支持大规模、网络促进的研究数据交换。

HHS 长期优先事项：

- 建立人工智能就绪的医学研究与发现数据集的治理、法律和分析框架作为公共资源。

1.6.2 推动可靠的人工智能开发以及合乎伦理道德且负责任的使用

随着人工智能在医学研究与发现领域的应用持续快速推进，其相关风险可能需要 HHS 密切关注，以确保其使用对全球患者安全、负责且具有影响力。为确保人工智能在医学研究与发现中的可靠及安全使用，HHS 可采取的关键行动主题包括：

1. 建立和传播证据以减轻生物安全、数据安全、隐私和数据收集风险
2. 制定明确的指南，用于医学研究与发现中安全和可信的人工智能使用以及联邦资源的分配和使用
3. 实现人工智能风险管理和透明度的安全和负责任组织治理

以下，HHS 更详细地讨论了每个行动主题的背景、迄今为止已采取的相应行动，以及为确保人工智能在医学研究与发现中可靠且安全使用的相关计划。

1. 建立和传播证据以减轻生物安全、数据安全、隐私和数据收集风险：

背景：

如 1.5.1 节所述，医学研究与发现中的人工智能可能被恶意用于制造生物安全和生物安全威胁（例如，潜在的新型病原体）。此外，机密、敏感或分类信息可能会通过人工智能模型训练和部署有意或无意地泄露，并且收集敏感患者数据可能需要患者的去标识化或授权，这两者都可能对收集用于医学研究与发现的具有统计效力的信息量构成挑战。

HIPAA 隐私规则对用于研究的患者信息的使用和披露有特定规定（注意，HIPAA 隐私规则对各种情况下受保护健康信息（PHI）的使用和披露有规定，在医疗保健提供章节中有更详细的概述），并且人工智能模型在遵守隐私保护方面存在独特考虑。潜在的患者担忧包括对其去标识化数据使用缺乏同意以及对其同意的个人数据使用方式缺乏透明度。人工智能使得利用各种数据集（包括公开可用的外部数据）重新识别信息变得更容易，这可能需要调整数据共享政策和实践，特别是对于不受 HIPAA 约束的实体。HHS 和联邦政府已经采取行动解决此问题，并且未来，HHS 将采取进一步行动继续保护医学研究与发现中人工智能使用的敏感信息。

HHS 迄今为止的行动（非详尽）：

- NIH 的数据管理和共享政策促进科学数据的共享，以部分帮助加速生物医学研

究发现，通过实现研究结果的验证、提供对高价值数据集的访问以及促进数据在未来研究中的再利用。它还强调了良好数据管理实践的重要性，并确立了期望，即最大限度地适当共享由 NIH 资助或进行的研究产生的科学数据，同时有合理的限制或例外。

○ **NIH 的数据管理和共享政策补充信息：保护参与者隐私在共享人类科学数据时**概述了研究人员在共享参与者数据时保护研究参与者隐私的原则、最佳实践和考虑要点。该框架不建立具有约束力的规则，而是提供了一个框架，用于共享可识别和去标识化数据以及在获得同意和未获得同意的情况下获得的数据。

- **实施总统执行办公室的国家生物防御战略**，该战略解释了美国政府将如何更有效地管理其活动，以评估、预防、保护、应对和从生物威胁中恢复，这可能隐含来自人工智能使用的威胁。
- **HHS 针对合成核酸供应商及用户的筛查框架指南**，阐述了其筛查框架指引内容。该指引为基因与基因组合成行业制定了基线标准，同时针对所有参与合成核酸供应、使用及转移的实体，在筛查订单与接收方以及记录保存方面，提出了最佳实践方案。此外，此指南旨在倡导最佳做法，以应对因合成核酸可能被滥用而引发的生物安全问题，防止其绕过现有监管控制并被用于实施非法行为。
- **实施总统执行办公室的核酸合成筛选框架**，该框架与 HHS 筛选框架指南一致且响应，并履行 2023 年关于安全、可靠和可信的人工智能开发和使用的行政命令的规定，该命令要求所有接受美国政府生命科学研究资助的研究人员仅从符合序列筛选最佳实践的公司采购合成遗传材料（88 FR 7519）。
- **HHS 的 HIPAA 隐私规则**确定了涵盖实体在研究目的下可以使用或披露受保护健康信息（PHI）的条件（45 CFR 第 160 部分和第 164 部分 A 和 E）。根据此隐私规则，涵盖实体在获得个人授权或在隐私规则规定的有限情况下可以使用和披露 PHI 用于研究。虽然隐私规则可能未明确讨论人工智能，但其保障措施无论人工智能是否在医学研究与发现中使用都适用。
- 由国家保护生物医学和行为研究人类受试者委员会编写的**贝尔蒙报告**是一份基本伦理原则和指南声明，应有助于解决围绕人类受试者研究进行的伦理问题，无论在研究和发现中使用何种技术，包括但不限于分析临床数据的医学研究与发现中的人工智能。

HHS 近期优先事项：

- 迭代监测和评估潜在的恶意使用，以不断完善与生物安全和数据泄露相关的指南和政策。
- 考虑根据**筛选框架指南**审查用于氨基酸和核酸序列筛选的预测方法。

- 促进公私合作过程以定义开放行业标准，以加速数据去标识化的隐私增强技术的可用性（例如，隐私保护记录链接（PPRL），差分隐私）。
- 评估潜在的技术解决方案，使开发人员和研究人员能够在沙盒环境中创建和使用模型，以防止数据泄露，从而实现医学研究与发现中人工智能使用的安全测试和进展。
- 探索在数据收集（包括数据质量，例如电子健康记录在某些临床环境中可能显示高质量与低质量结果）中利用人工智能的机会和风险。
- 探索潜在的数据使用授权途径，使患者数据能够在迭代和可能的多用途人工智能模型中使用，同时保持与 HHS 值、法规和政策一致的保护。
- 探索为评估同态加密和数据安全提供资源，以实现数据的联合而不允许查看数据链接，从而确保医学研究与发现中人工智能的安全使用。
- 探索保护医学研究与发现中使用的人工智能模型和敏感健康数据免受对抗攻击的方法。
- 探索开发机制以防止和减少医学研究与发现中预测分析工具滥用造成的危害。
- 提供关于在患者、参与者、基因组和受控访问数据上训练模型的指南，因为存在高数据泄露风险以及隐私和保密问题。考虑征求社区意见以告知这些指南。
- 探索保护敏感健康信息的数据共享协议。

HHS 长期优先事项：

- 考虑潜在的政策解决方案或指南，使医学研究与发现能够在受控访问环境之外利用人工智能，同时最大限度地减少数据泄露风险。
- 提供关于研究中人工智能的特殊考虑的政策清晰度和（或）指南，包括专门为研究开发的人工智能的定义、人工智能模型用于研究的可用性、用于研究的患者数据的重新识别风险以及研究背景下人工智能的隐私和安全影响。
- 评估与 STLTs 合作的潜在途径，以使患者能够授权其数据在医学研究与发现中的使用，以增强医学研究与发现中的多样性和代表性，同时设计长期解决方案以加速和扩大安全数据收集和使用。
- 考虑潜在的技术或政策解决方案，最大限度地减少患者数据收集的障碍，同时维护数据安全并最大限度地减少未经授权的使用。

2. 制定明确的指南，用于医学研究与发现中安全和可信的人工智能使用以及联邦资源的分配和使用：

背景：

建立和促进可信人工智能对于医学研究与发现中负责任地使用人工智能至关重要。

为医学研究与发现中使用的人工智能模型开发证据并传播关于透明度和其他伦理、法律和社会影响（ELSI）的指南和监管期望，包括那些利用联邦资源的模型，可能会导致该领域更安全和更可信的人工智能使用。HHS 已经采取措施解决这一挑战，并将在未来继续建立保障措施。

HHS 迄今为止的行动（非详尽）：

- HHS 政策制定者已经建立了一个**监管框架，称为共同规则，以指导生物医学研究**。这个框架将继续支持整个研究生命周期中人工智能的伦理和负责任使用。附录 B 包括详细说明在研究中开发和使用人工智能之前、期间和之后应如何考虑这些法规、政策和最佳实践的特定网页。该政策框架的主要原则包括：
 - **保护人类受试者研究参与者**，旨在保护研究参与者的权利、安全和福利。
 - **健康信息隐私政策、法规和最佳实践**有助于保护研究中使用的健康数据的隐私和安全，从而促进对医疗保健研究活动的信任。
 - **生物安全和生物安全监督**继续适用于生物医学研究中人工智能的开发或使用。
 - 关于**公众访问研究产品以及数据管理和共享**的政策和指南，旨在最大限度地负责任和适当地共享和管理研究产品，同时确保研究人员考虑如何保护人类研究参与者的隐私、权利和机密性。负责任和适当的共享和管理不仅指人类数据保护，还包括其他相关法律、法规和政策，这些政策限制披露并对共享施加限制。
 - 与知识产权和软件共享相关的**许可、知识产权和技术转让政策**和资源，以补充数据共享并界定研究人员的权利。
- **NIH 的人工智能在研究政策考虑和指南**详细阐述了一套健全的政策和实践体系，用以指导生物医学和行为研究生态系统中的利益相关者。NIH 的政策框架旨在负责任地指导和管理科学进步和新兴技术，包括在研究中开发和使用人工智能技术。所讨论的政策、最佳实践和法规反映了这个框架，并且在研究中开发和使用人工智能之前、期间和之后都应予以考虑。它不是可能适用于任何 NIH 支持的研究项目的所有政策和要求的详尽列表，但它可以指导研究社区关于隐私、知识产权、数据管理、参与者保护等方面。

HHS 近期优先事项：

- 协调中游（例如，NIH）和下游（例如，FDA）医学研究与发现机构之间的信息共享，并在可能的情况下协助寻求监管授权的开发人员。
- 探索开发一个共同的期望框架，用于解决或提供关于使用人工智能进行医学研究与发现中，研究人员如何处理 ELSI 以进入临床试验和潜在监管批准的透明

度。

- 考虑支持指南和教育工具，以帮助人工智能开发人员在创建用于医学研究与发现的人工智能技术时朝着安全、安全和信任的方向努力。
- 探索针对研究资源、培训和研讨会的目标，以进一步研究医学研究与发现中人工智能的 ELSI，包括可解释人工智能。
- 创建实践社区（例如，沙盒）的机会，以在内部以降低成本评估医学研究与发现中人工智能技术的 ELSI。

HHS 长期优先事项：

- 根据需要实施更新和（或）新政策，以确保在内部（例如，通过 HHS 和（或）HHS 资助或合同接受者）和外部（例如，在工业和（或）学术界）医学研究与发现中负责任地使用人工智能，包括可能对医学研究与发现中的人工智能风险进行分层。
- 随着该领域的不断发展，持续将资源配置的重点放在探索如何构建证据体系上，以评估人工智能在医学研究与发现中的伦理、法律和社会影响（ELSI）。
- 持续监测医学研究与发现中人工智能的进展，定期更新和修订政策和（或）指南，以进一步明确与后期监管批准过程、ELSI 以及药物和生物产品批准和设备营销授权要求相关的人工智能使用。

3. 实现人工智能风险管理和透明度的安全和负责任组织治理：

背景：

人工智能的可靠使用依赖于对模型性能与特性的保证，以及决定人工智能在实际中如何应用的实施流程及相关工作流程。目前已有大量关于涵盖人工智能应用的负责任研究实践的政策指导。然而，针对人工智能在医学研究与发现中的应用，若缺乏专门的风险管理政策，无论技术质量如何，都可能导致人工智能表现不佳。

此外，社区能够协助识别与其居民相关的风险，并就透明度目标达成共识，这有助于降低人们对自身数据使用方式丧失信任的风险。目前，用于呈现医学研究与发现中所使用人工智能模型特性的标准化方法有限，难以让用户和监管机构更好地了解特定人工智能模型的潜在隐患。HHS 已通过资助和研究此类技术来应对这一挑战。HHS 将继续分享指南、制定政策，并探索有助于实现这些目标的资源配置活动。

HHS 迄今为止的行动（非详尽）：

- **ARPA - H 的人工智能性能和可靠性评估用于连续修改和可用性项目（PRECISE - AI）** 资助研究，旨在开发一种技术，该技术能够检测在现实临床护理场景中使用的的人工智能何时与基础训练数据出现偏差，并且重要的是，能

够自动纠正这种偏差。

- **退伍军人事务部 (VA) 和 FDA 即将合作的虚拟健康人工智能实验室**将在虚拟实验室环境中测试医疗人工智能应用程序，以确保它们对退伍军人和患者有效、安全，并遵守可信人工智能原则。
- **HHS 的可信人工智能框架**描述了可以采取哪些方法来解决医疗保健中与人工智能相关的许多伦理和其他挑战，包括那些可能适用于医学研究与发现的挑战。虽然它不是官方政策，但它可以澄清 HHS 如何应对与人工智能使用相关的这些挑战。
- **AHRQ 的数字医疗公平框架**指导用户有意考虑涉及数字技术的医疗保健解决方案中的公平性，并评估这些解决方案在每个数字医疗生命周期阶段是否公平。

HHS 近期优先事项：

- 探索在数据收集（包括数据质量，例如电子健康记录展示高质量与低质量结果）中利用人工智能的机会和风险。
- 探索合成数据风险管理的技术或政策解决方案，以减少随着分析和后续合成数据生成迭代而可能发生的合成数据退化。
- 制定与更广泛的 HHS 质量保证政策和计划一致的用于研究中使用的的人工智能的质量保证计划，包括所有规划、开发和发布的数字无障碍性。
- 探索减轻滥用的策略并确定评估当前人工智能模型、数据集和研究结果风险的方法。
- 与其他联邦机构协商，更新和完善联邦资助研究活动的风险管理指南，以主动识别、评估和减轻与研究中使用的的人工智能相关的风险。
- 定义、优先考虑并传播用于测试、评估、验证和验证医学研究与发现中使用的算法的框架。
- 探索鼓励向价值链中其数据可能有助于开创性研究的利益相关者透明地使用人工智能模型和个人数据的机会，包括相关风险。
- 对那些在人工智能相关主题方面技能、经验和知识相对欠缺的研究人员以及普通公众进行培训，确保他们了解人工智能在医学研究与发现中应用时可能存在的两用性风险及其他风险。
- 探索利用人工智能动态评估医学研究与发现中人工智能风险的潜在应用，鉴于模型的动态性质和当前静态的风险管理框架。

HHS 长期优先事项：

- 探索隐私增强技术及其在 HHS 支持和开展的涉及人工智能的研究中的潜在

应用。

- 与行业合作开发“研究模型卡”框架，用于标准化表示医学研究与发现中使用的人工智能模型的特征，包括（1）设计目的，（2）关键开发输入，（3）关键模型输出，（4）外部验证过程和结果；以及（5）生命周期管理计划和过程。

1.6.3 普及人工智能技术和资源

人工智能方法有潜力为研究人员“创造公平竞争环境”，帮助在广泛、丰富、多模态和复杂的数据集中识别以前无法检测到的模式，就像 CRISPR 使基因编辑在全球范围内广泛可用一样。然而，让更多研究人员能够接触（相关技术），以及将其应用于更多尚未得到充分投资的技术领域，这些情况不会自发出现；联邦政府的引导、激励措施和政策，在确保人工智能技术被用于那些市场自身可能无法充分或迅速达成的目标方面，发挥着关键作用（见目标 1：“促进健康人工智能创新与应用”行动主题 1：“扩大医学研究与发现中人工智能在疾病领域和价值链步骤中的使用范围”以获取更多信息）。虽然创新已经超出了实验室范围，但一些利益相关者可能仍然缺乏参与人工智能的资源，关键行动主题包括：

1. 促进有意的公众参与和公私合作行动，以增强所有利益相关者之间的最佳实践共享
2. 增加所有人对合理管理的人工智能就绪数据、模型和算法以及工具和基础设施的可及性

以下，HHS 更详细地讨论了每个行动主题的背景、迄今为止的相应行动以及确保公平获取人工智能技术和资源的计划。

1. 促进有意的公众参与和公私合作行动，以增强所有利益相关者之间的最佳实践共享：

背景：

增加利益相关者之间的合作伙伴关系（例如，行业、STLTs、学术界和公众）以及在整个创新管道中的有意公众参与，可以通过在群体之间共享想法、方法、最佳实践、示例应用和关键风险缓解措施来增强人工智能在医学研究与发现中公平使用的潜力。HHS 已经开始召集利益相关者，并将继续采取行动应对这一挑战。

HHS 迄今为止的行动（非详尽）：

- **NIH 的 AIM - AHEAD 计划**寻求与代表性不足的社区建立伙伴关系，以在行为和生物医学研究中开发和使用人工智能，建立网络以解决健康差异。该计划通过改善社区参与、领导力和研究奖学金（特别是在服务不足的社区）的项目来推动研究和指导，并促进研究中人工智能的基础设施发展。
- NIH、NSF、NCSES 和 DOE 的**国家人工智能研究资源 (NAIRR) 试点**是一个跨机构合作项目，致力于改善研究中的人工智能，包括与人类健康相关的主题。它利用大型真实世界数据 (RWD) 集来（1）构建合成数据生成器工具包和框

架以评估使用此类数据进行证据构建的隐私风险和效用，以及（2）将医学成像数据与临床记录链接，以建立多模态人工智能开发的能力。

- **NIH 的“我们所有人”研究计划**是一个全国性的参与者伙伴和研究人员网络，旨在确保来自所有背景的人都能参与研究。参与者慷慨分享信息，为数千项研究提供动力，以更好地理解健康和疾病，实现更量身定制和公平的护理方法，并创造利用人工智能推进精准医学的新机会。
- **HHS 还在开发挑战（即创新竞赛），举办研讨会（例如，“人工智能在人类研究中的演变格局”），并与咨询委员会合作，征求公众对促进数据访问、组合和分析的工具（例如，人工智能、云计算）的意见。**

HHS 近期优先事项：

- 促进和便利人工智能开发者与 NIH 资助的研究人员之间的公私合作伙伴关系的法律途径（例如，通过美国国立卫生研究院基金会）。
- 制定愿景和框架，将公众声音纳入所有阶段和类型的临床研究。
- 探索以易于理解的形式让公众参与和教育关于医学研究与发现中人工智能的益处、风险和潜在用途的机会，以建立信任并促进公平使用。
- 继续让利益相关者（见图表 3），包括公众和参与者，参与医学研究与发现管道，以收集他们对人工智能应用的想法。
- 扩大合作机会和实施提高 NIH 支持的数据的人工智能就绪性的倡议。
- 促进公私合作，促进 NIH 资助的研究机构与服务不足或代表性不足的机构之间的人工智能知识和技术共享。

HHS 长期优先事项：

- 探索增加对多机构研究合作的资源投入，特别是那些嵌入生物伦理学家和开发者的合作。
- 提供安全沙盒和基础设施，鼓励对医学发现中人工智能的开发和使用进行合作研究，前提是它们确保信息和通信技术（ICT）的开发符合 HHS 数字无障碍指南。
- 促进社区参与，这将培育、发展和维持研究人员与公众成员之间的长期关系，可用于共同创作。可能需要新的授权来调查利益相关者（包括通过人工智能，考虑或获得《减少文书工作法》的豁免）。可能需要新的政策来负责任地监管此类伙伴关系。

2. 增加所有人对合理管理的人工智能就绪数据、模型和算法以及工具和基础设施的可及性：

背景：

有效和高效地利用人工智能需要财务、技术和人力资源。虽然不是商品，但通用人工智能技术（例如，大型语言模型）广泛可用，并可能“提高行业整体能力的下限”。更多样化的研究人员和用例在医学研究与发现中应用这些技术的潜力可能会受到资源可用性的阻碍，这可能会加剧已经普遍存在的“数字鸿沟”。HHS 已经使数据和工具更易于获取，并计划继续迭代这些活动。

HHS 迄今为止的行动（非详尽）：

- **NIH 的科学技术研究基础设施用于发现、实验和可持续性 (STRIDES) 倡议**为 HHS 资助的行为和生物医学研究人员提供折扣访问商业云服务，包括人工智能应用。STRIDES 已经为这些研究人员节省了约 1.2 亿美元的成本，他们还可以访问相关的“云实验室”，这是一个沙盒，带有相关教程和数据，研究人员可以在其中免费试用这些技术。
- **NIH 的数据管理和共享政策**要求研究人员前瞻性地规划最大限度地适当共享“科学数据”（即具有足够质量以验证和复制研究发现的数据）并遵守 NIH 批准的计划。该政策的补充信息帮助研究人员选择数据存储库、为数据管理和共享进行预算并保护人类研究参与者数据。
- **NIH 的“我们所有人”研究计划**，在上述“促进有意的公众参与和公私合作行动，以增强所有利益相关者之间的最佳实践共享”行动主题中也有提及，正在构建一个多样化的数据库，可为关于各种健康状况的数千项研究提供信息。该计划创建了有史以来最大、最多样化和最广泛可访问的健康研究数据集之一。研究人员可用的数据包括基因组数据、调查问卷回复、身体测量、电子健康记录信息和可穿戴设备数据。该计划的基于云的平台设计鼓励跨机构合作，使研究人员能够利用人工智能和相关工具并扩大他们对许多健康状况的理解。
- **ARPA - H 的生物医学数据 fabric 工具箱**与 NIH 合作，旨在通过以下方式更容易地连接来自数千个来源的生物医学研究数据：（1）降低以计算机可读形式进行高保真、及时数据收集的障碍，（2）为大规模多源数据分析做准备，（3）推进直观的数据探索，（4）在维护隐私和安全措施的同时改善利益相关者的访问，以及（5）确保生物医学数据 fabric 工具在不同疾病类型中的通用性。这些数据必须是可查找、可访问、可互操作和可重用的。
- **NIH 的通用存储库生态系统倡议**支持七个通用存储库，它们共同努力建立一致的元数据、开发数据共享和重用的用例，并培训和教育研究人员如何共享和重用数据，包括用于人工智能的开发和使用。

HHS 近期优先事项：

- 探索针对研究资源、培训和研讨会的目标，以“扩大”有能力进行人工智能研究的机构基础，可能重点关注数据基础设施。

- 探索为机会提供资源，继续支持资源较少的机构获得对人工智能在医学研究与发现中使用至关重要的基础设施（例如，存储、计算、模型）。
- 扩大像 NAIRR、GREI 和 ScHARe 这样的资源的可用性和能力。
- 评估 STRIDES 计划的扩展，以包括人工智能工具和模型。
- 扩大联邦数据倡议的可用性、能力以及知识和工具/技术共享。
- 开发作为公共资源的授权利益相关者可用于医学研究与发现的联合、链接、集中式人工智能就绪数据存储库。
- 继续开发可公开利用以通过人工智能生成见解来指导医学研究与发现的数据平台。

HHS 长期优先事项：

- 提高协助研究人员根据不断变化的研究公共访问情况完善数据管理和共享标准的能力。
- 构建一个内部数据库，跟踪医学研究与发现中的合规性、公众评论和其他人工智能可及性问题。

1.6.4 培养人工智能赋能的劳动力和组织文化

如果没有足够的人工智能专家来实现医学研究与发现的大规模创新，广泛的使用和应用可能是不可行的。为此，HHS 计划在内部和外部刺激劳动力发展，以授权在医学研究与发现价值链中持续进行负责任、安全的人工智能创新。该领域当前的行动主题包括：

1. 改善医学研究与发现中人工智能治理和管理方面的培训
2. 在医学研究与发现中培养和保留强大的人工智能人才库

以下描述了 HHS 当前的行动和未来在医学研究与发现中创建人工智能赋能的劳动力和组织文化的目标。

1. 改善医学研究与发现中人工智能治理和管理方面的培训：

背景：

大多数参与人工智能的人将负责管理和使用这些技术，而不是开发它们。确保医学研究与发现企业从人工智能中获得最大收益将需要关注技术，也许更重要的是，关注其实施、工作流程集成和生命周期管理。培养医学研究与发现领域的专业人才，使其能够负责任地管理和使用此类技术，这对于利用人工智能推动该行业发展也至关重要。HHS 已着手应对这一挑战，并将进一步加大工作力度，以弥合这一差距，助力该行业发展。

HHS 迄今为止的行动（非详尽）：

- FDA 的博客文章“安全、有效的人工智能辅助医疗服务的生命周期管理方法”提供了一种开发、验证和维护医疗设备人工智能模型持续治理的潜在方法概述，以确保其安全性和有效性。

HHS 近期优先事项：

- 探索针对资源、培训和研讨会的目标，包括研究和技术中人工智能技术的治理、管理和使用。
- 考虑支持关于医学研究与发现中人工智能技术的治理、生命周期管理和工作流程集成的指南或最佳实践。

HHS 长期优先事项：

- 根据需要迭代修订和发布指南或培训计划的更新。

2. 在医学研究与发现中培养和保留强大的人工智能人才库：

背景：

为了利用人工智能在医学研究与发现中的潜力，生态系统可能需要一个强大而多样化的劳动力供应链，能够将模型和算法集成到他们的研究中。随着多模态模型变得越来越强大并可能自动化科学工作流程的许多方面（从观察和假设开发到数据分析和手稿编写），不同类型的人工智能可能会改变有效医学研究与发现劳动力所需的技能组合和角色，在所有阶段都需要人类输入和评估。来自所有背景的研究人员可能需要基础知识来安全、负责任和有效地开发和应用人工智能。此外，如果没有明确的激励措施，跨学科专家可能会继续流向技术部门，在专注于医学研究与发现的非营利、学术和政府实验室中留下人才缺口。HHS 已经采取行动应对这一挑战，并计划继续探索机会。

HHS 迄今为止的行动（非详尽）：

- **NIH 的 AIM - AHEAD 计划**建立了一个强大的指导网络，以在美国培养医学研究与发现中的人工智能人才。
- **NIH 的 DATA 国家服务学者计划**聘请数据科学专业人员到 NIH，以提高研究效率、创新研究、工具开发和分析能力。
- **NIH 的信息科学、人工智能和生物医学科学界面的劳动力发展行政补充支持**在信息科学、人工智能和生物医学科学界面开发和实施课程或培训活动，以培养使生物医学数据可查找、可访问、可互操作和可重用以及为人工智能做好准备所需的能力和技能。
- **国家医学图书馆（NLM）的基于大学的生物医学信息学和数据科学研究培训计划**支持在美国的研究生和博士后教育机构进行生物医学信息学和数据科学研究培训。

- **NLM 的短期研究教育体验项目，吸引优秀学生投身生物医学信息学/数据科学事业并提升多样性：**旨在支持一系列教育活动，鼓励优秀的本科生和硕士生，包括那些来自在生物医学和行为科学领域代表性不足群体的学生，继续在生物医学信息学和数据科学领域接受深造并从事相关职业。NLM 希望通过该项目培养一批多样化的科学家，使其有能力引领生物医学信息学和数据科学研究。
- **NLM 的数据科学和信息学（DSI）学者计划**是一个为期 8 - 12 周的夏季实习项目，实习生在生物科学计算研究项目中贡献他们的技能和观点。DSI 学者在协作研究环境中获得宝贵经验，同时与研究导师进行一对一培训。

HHS 近期优先事项：

- 优先考虑并探索为证据建设提供资源，以评估人工智能劳动力发展努力并最大化 HHS 支出的效果。
- 加强并扩大针对研究人员的培训，内容聚焦于为医学研究与发现开发负责任的人工智能工具，其中涵盖将人工智能相关课程融入生物医学研究培训课程体系的最佳实践方法。
- 整合生物安全资源或培训内容，向刚接触人工智能应用的研究人员分享。
- 为医疗服务提供者制定关于在医学研究与发现中使用人工智能的教育和培训计划，内容包括如何安全地使用和收集患者数据，以推动进一步创新。
- 评估 **NIH 的人工智能推动健康研究与发现加速创新（AIM - AHEAD）**计划的拓展情况，包括针对医学研究与发现领域人工智能专业人才的招募与培训工作。

HHS 长期优先事项：

- 探索拓展资源配置机制，重点支持在生物医学研究研究生培养中，人工智能的开发与应用。
- 探索在数据科学与人工智能表现卓越的美国各研究中心提供资源支持，为 HHS 资助的研究人员提供补贴培训与服务。
- 推动以社区为导向的培训，提升在提示工程、红队测试和水印技术方面的技能，以便在保持科学严谨性和促进公平性的同时，最大限度地发挥人工智能的效用。

1.7 结论

在人工智能驱动的医学研究与发现领域，促进创新的同时管控风险，对于推进美国的卫生与公众服务事业至关重要。HHS 深知，人工智能在提升研究成果、加速医疗产品研发以及改善患者护理方面潜力巨大；然而，必须在这些益处与偏见、数据滥用、生物安全及其他相关风险之间寻求平衡。HHS 在这一领域具有独特优势，能够发挥关键作用。

HHS 的行动计划涵盖探索资源配置、公众教育以及劳动力发展等举措，旨在应对当前医学研究与发现中使用人工智能所面临的挑战，推动其安全、负责任地应用。通过这些努力，HHS 能够刺激经济增长、创造高技能就业岗位，而最重要的是，保障全体美国民众乃至全球个体的健康与福祉。通过发挥战略领导力并与价值链上的各利益相关方开展协作，HHS 能够引导人工智能在医学研究与发现中得到负责任的整合，助力确保在实现创新益处的同时降低相关风险。随着技术和应用场景的不断变化，HHS 致力于不断完善其在医学研究与发现领域的人工智能战略，以最大程度地推动医学研究与发现的进步。

2. 医疗产品开发、安全与有效性

2.1 引言与背景

医疗产品，包括药物、生物制品和医疗设备（包括一些基于软件的行为干预措施），在促进健康方面发挥着至关重要的作用。随着人工智能变得越来越先进，它有可能通过增强医疗从业者的能力并在从临床试验到制造和安全监测的整个生命周期中支持产品开发，进一步改善患者护理。人工智能技术在医疗产品领域的快速发展使 HHS 处于关键地位。HHS 可以推动有效技术的成功使用和扩大规模，同时最大限度地减少医疗产品在整个生命周期中与潜在风险和危害相关的问题。

本章将重点关注医疗产品本身以及从临床试验到监管审查、制造和安全监测的医疗产品生命周期步骤。有关医疗产品的研究与发现以及可在生物医学中应用的人工智能技术的研究与发现的更多信息，请参阅“医学研究与发现”章节。

人工智能在医疗器械中的作用与其他医疗产品有所不同。在药品和生物制品方面，人工智能通常有助于生成信息或数据，以支持产品从研发、生产到上市后监督与监测等整个开发生命周期中的决策。在医疗器械领域，人工智能可能发挥三种作用：用于器械的研发或维护；作为独立产品，无需成为传统硬件设备的一部分，就能实现一项或多项器械功能（如诊断、治愈、缓解、治疗或预防疾病）；或者作为器械的组成部分，甚至是核心部分。

在美国，这些产品的上市授权监管审查由法定和监管框架管理，有助于确保医疗产品在其预期用途上是安全和有效的。在整个产品生命周期中，FDA 在产品上市前审查有关产品的数据和信息，在产品上市后进行监测，并监督产品推广和医疗产品质量。

截至 2024 年 8 月，FDA 已授权约 1000 种人工智能赋能的医疗设备，并且 FDA 已收到超过 550 份包含人工智能组件的药物和生物制品提交。NIH 也在推动医疗产品的开发以增加获得更好护理的机会方面发挥着关键作用。虽然临床开发资金可以来自多种渠道，但 NIH 每年仅在临床试验方面的投资就达到 30 亿美元，使其成为美国最大的临床试验联邦资助者。政府对医疗产品的监管，致力于在保障安全性与有效性和促进创新之间维持平衡，当人工智能应用于医疗产品或贯穿医疗产品的生命周期时，亦当如此。

2.1.1 行动计划摘要

本章后面，HHS 将阐述为推进该部门在负责任使用人工智能方面的四个目标而提出的行动建议。以下是每个目标下行动主题的概述。有关行动建议的完整详细内容，请参阅 2.6 节“行动计划”。

行动支持的关键目标	拟议行动主题（非详尽，详见 2.6 行动计划）
1. 促进健康人工智能创新与应用	-明确医疗产品的监管机制 -明确支付模式 -促进公私伙伴关系和政府间合作，以快速开发和共享知识
2. 推动可靠的人工智能开发以及合乎伦理道德与负责任的使用	-完善监管框架，以应对医疗设备中自适应人工智能技术 -促进人工智能部署的公平性，加强安全和负责任的使用 -解决当前设备监管机构管辖范围之外的人工智能赋能软件问题 -建立用于健康领域人工智能质量保证的私营或公共机制
3. 普及人工智能技术和资源	-通过公众参与实现协作开发 -统一研究和医疗服务领域的标准和信息共享机制
4. 培养人工智能赋能的劳动力和组织文化	-加强医疗产品中人工智能治理和管理方面的培训 -培养并留住医疗产品相关的人工智能人才

2.2 参与医疗产品开发、安全与有效性的利益相关者

众多利益相关者会参与到医疗产品及其开发过程中的人工智能应用，涵盖从患者、医疗服务提供者，到医疗产品开发者、经销商、供应商、支付方、研究人员等诸多群体。本章末尾的“行动计划”部分包含了吸引这些利益相关者参与的方法，以便在推进创新的同时降低风险。以下是一个示意性图表，展示了利益相关者之间的示例流程，以及一份项目列表，包含了参与医疗产品开发、安全性和有效性的利益相关者的更多详细信息。请注意，该图表和列表并未涵盖所有可能的利益相关者角色及互动。如需获取有关监管指导和权限的更多详细信息，请参考 HHS 的其他文件。

图 5：参与医疗产品开发、安全 and 有效性的利益相关者

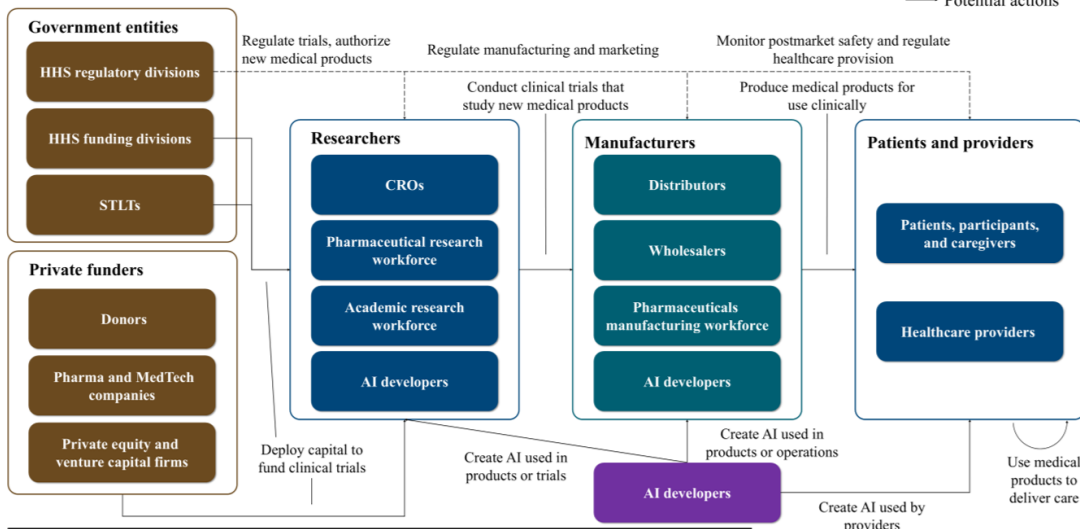
非详尽示例

NON-EXHAUSTIVE | ILLUSTRATIVE

Below is an example of the flow of authorities and actions in medical product development, safety, and effectiveness. It does not capture all the permutations and intricacies of stakeholder roles and interactions, and entities can play multiple roles.

..... Potential authorities

—— Potential actions



Please see information on official FDA, OCR, ASTP/ONC, NIH, and other HHS websites for more detailed information on regulatory authorities in medical product development, safety, and effectiveness.

利益相关者包括但不限于：

- **HHS 运营部门（非详尽列举）：**

- **FDA：**有助于确保人用和兽用药物、生物制品和医疗器械对于其预期用途是安全有效的，以及发射辐射的电子产品是安全的。随着人工智能在医疗产品及其开发、制造运营和使用中变得更加突出，FDA 将继续监管和支持利益相关者。
- **NIH：**支持美国国内外的生物医学和行为研究，在其实验室和诊所开展研究，培训有前途的年轻研究人员，并促进生物医学知识的收集和共享，这些研究越来越多地涉及与医疗产品和生命周期相关的人工智能。
- **CDC：**在 FDA 批准疫苗后制定使用建议，在临床使用后持续监测疫苗安全性，并报告不良反应（例如通过疫苗不良事件报告系统）²⁰⁹。
- **AHRQ：**支持对医疗器械进行的干预措施（如以患者为中心的临床决策支持）的研究，并专注于提高所有美国人医疗保健的质量、安全、效率和有效性。

- **其他联邦机构：**HHS 还与许多其他联邦部门密切合作，如国家科学基金会（NSF）和能源部（DOE）。

- **患者、参与者和护理人员（包括居民和社区）：**使用利用人工智能开发或包含人工智能的药物、生物制品或医疗器械。如今，有能力的患者也可能利用人工智能更好地了解自己的个人健康状况并为自己的护理进行倡导。

- **制药和医疗技术与制造公司：**设计、开发和生产用于提供医疗保健服务的商业药物、生物制品或医疗器械，包括将人工智能集成到临床试验和产品设计制造中的研究人员和主题专家。他们是临床试验和医疗产品制造中人工智能的主要用户之一。这些公司还使用人工智能支持药物警戒活动。

- **医疗服务提供者和支付者：**利用医疗产品并为临床开发工作提供临床视角（例如医院、诊所、医疗专业人员）或决定哪些技术属于其支付机制（例如支付者）。

此外，医疗服务提供者可以是人工智能使用中的“人在回路”，这包括医疗产品生命周期的部分环节。临床环境中人工智能的使用在“医疗保健交付”章节中有更详细的阐述，因为制造商预期和 FDA 授权的医疗产品使用可用于为特定目的提供医疗保健，而不改变其设备、药物或生物制品的状态。

- **STLTs**: 在联邦政府之外发挥监督和资助作用。FDA 对医疗产品进行监管监督，而 STLTs 对医疗实践和医疗保健交付的不同组成部分拥有管辖权。
- **学术、非营利和其他研究人员**: 为生物学知识的前沿发展提供证据，包括设计和生成用于临床应用的医疗器械的工程师，以及开发人工智能、将人工智能应用于临床试验工作流程和 / 或将人工智能集成到产品开发生命周期中的主题专家。他们是医疗产品开发中人工智能的主要用户之一。
- **合同研究组织 (CROs)**: 提供外包研究服务，并将人工智能开发或集成到其临床试验工作流程中。作为第三方，CROs 在处理其他组织的敏感数据时应特别关注安全和隐私问题。这些公司还使用人工智能支持可能由制药商外包的药物警戒活动。
- **分销商和批发商**: 促进医疗产品（可能包括或利用人工智能进行研究和开发）向医疗服务提供者的分销。
- **捐赠者和私人资助者**: 支持产品开发和扩大规模的资金。他们包括非营利性捐赠者，如基金会，以及营利性资助者，如私募股权、风险投资和其他资助组织。这些组织还可能支持对人工智能技术或与其他利益相关者的其他投资。
- **人工智能开发者**: 构建可用于医疗产品或整个医疗产品生命周期的人工智能工具、模型和平台。

2.3 人工智能在医疗产品开发、安全和有效性方面的应用机会

如果人工智能在医疗产品开发、运营和安全管理中得到成功且可靠的使用和推广，以及被纳入医疗产品本身，可能会通过多种方式改善整体治疗效果以及流程的可及性和效率，例如：

1. 提高临床试验效率，加速获得安全有效的医疗产品：在临床试验中利用人工智能可能有助于预测参与者的不良反应风险，生成监管提交文件和研究手册的初始内容，并将文档翻译成其他语言。此外，虽然有不依赖人工智能实现以患者为中心的方法，但朝着这个目标使用人工智能可能会降低候选者流失的可能性。利用人工智能进行分析可以加速临床试验过程的另一个核心部分。总之，这些以及其他在临床试验中使用人工智能的方式可以使患者更快地获得医疗产品（有关迄今为止临床试验中人工智能应用的更多详细信息，请参阅下文趋势（A）（1））。

2. 提高临床试验对医疗产品使用者的代表性：如今，多达“86% 的临床试验在规定时间内未达到招募目标”，这可能导致医疗干预效果不佳，在制药实践中可能导致健康公平性较差，并可能造成数十亿美元的经济损失。在临床试验策略中适当利用人工智能分析患者和其他人口统计数据、选择试验地点并识别代表目标人群的潜在候选者，

有可能帮助招募更具代表性的临床试验人群。这可以为提交给 FDA 进行监管批准或市场授权的信息提供支持。在临床试验策略中利用人工智能可以更好地服务于历史上代表性不足的人群。

3. 作为医疗产品的一部分、医疗产品本身或用于开发医疗产品：人工智能可以作为医疗产品的一部分或用于开发安全有效的医疗产品。特别是，支持人工智能的医疗器械，如非处方助听器，有可能被患者、医疗服务提供者和其他终端用户用于帮助加强护理和改善结果。此外，人工智能支持从临床使用期间收集的数据中学习的能力，这有助于随着时间的推移提高医疗产品的准确性和性能，可能导致提高准确性和监测（例如降低误诊率、提高早期检测不良反应的能力）。同样，人工智能可用于开发药物和生物制品（例如识别靶点、评估生物标志物和终点），如在“医学研究与发现”章节中所讨论的。

4. 加强供应链、制造和其他运营，确保并扩大可及性：近年来，医疗产品供应短缺影响了患者及时获得对其健康至关重要的护理的能力。例如，截至 2024 年 10 月，有超过 100 种药物短缺，从静脉输液到处方兴奋剂不等。同样，当对特定医疗产品的需求激增时，通过迅速增加供应来扩大可及性可能不是一个快速的过程。人工智能可以合理化和简化供应链管理和制造流程，包括分析生产计划、预测需求、估计潜在中断的影响和优化库存。通过负责任地将人工智能应用于其运营中，医疗产品制造商、分销商和其他相关方可以缓解短缺问题，保障获得护理的机会，并在需求激增时为更多患者做好准备。

5. 加强药物警戒和上市后监测与监控：监测医疗产品对于管理其安全有效使用至关重要。如今，数据收集和分析已经利用电子健康记录（EHRs）、行政索赔和其他临床数据源来整理大量产品安全数据（例如 FDA 的不良事件报告系统 [FAERS] 和 FDA 的哨点计划）。一些安全监测活动涉及调查和社交媒体监测，这可能需要大量资源和时间。利用人工智能收集和 / 或分析不良事件报告、抓取的社交媒体数据或调查数据的大型数据集，可以快速识别潜在安全问题并加速采取行动保护患者的时间线。此外，这些数据和分析可用于更好地了解医疗产品使用的结果并获得新的见解以增强人类健康，包括特定医疗产品最适合服务的患者类型。然而，需要注意的是，由于可能有大量临床数据，可能会产生更多噪声，因此从数据中解析关键信号至关重要。

2.4 人工智能在医疗产品开发、安全和有效性方面的趋势

利益相关者已经开始在医疗产品及其开发中沿着两个总体趋势利用人工智能：

- A. 在医疗产品的开发和持续运营中利用人工智能
- B. 将人工智能嵌入产品本身或作为独立产品

以下讨论了迄今为止在（A）和（B）方面的一些非详尽的使用示例。

A. 医疗产品开发和运营中的人工智能

1. 与药物和生物制品开发相关的人工智能应用正在增加：在一系列治疗领域（TAs）中，人工智能在药物和生物制品开发生命周期中的使用越来越多。事实上，FDA 看到在过去几年中使用人工智能组件的药物和生物制品申请数量大幅增加，从 2018 年的仅 3 个增加到 2021 年的 132 个。这些提交涵盖了从临床研究到上市后监测和先进制药制造的药物和生物制品开发领域。最近 FDA 提交文件中看到的用例集中在一系列主题上，包括但不限于终点和生物标志物评估、异常检测、成像、视频和语音分析、患者风险分层和管理、剂量选择和优化以及临床试验期间的依从性。其他用例涵盖了临床试验中一些最耗时的方面（例如试验地点选择和候选者招募），并有助于预测拟议试验设计的成功或失败。人工智能还被用于通过选择参与者和减少错误来减少与随机对照试验相关的时间并提高其质量。

2. 验证健康人工智能可信度的方法各异且应用不一致：在健康领域（包括医疗产品的开发和运营）中使用人工智能需要进行验证，以确保其导致安全有效的医疗产品、决策和行动。如今，有许多人工智能验证方法，一般来说，它们侧重于在狭窄且高度受控的条件下易于测量的定量性能指标，很少使用真实患者数据。人工智能能够轻松部署到广泛且不断扩展的医疗用例中，这促使人们需要建立国家认可的标准和机制来确保人工智能系统的质量。

B. 产品内部或作为产品的人工智能

1. 支持人工智能的医疗器械应用正在扩展，重点在放射学领域：在医疗器械领域，人工智能迅速发展，涵盖了医疗产品生态系统中的新应用。截至 2024 年 8 月，FDA 已审查并授权约 1000 种支持人工智能的医疗器械在美国上市，其中 2023 年有 171 种，2024 年有 258 种，这表明在过去两年中授权的支持人工智能的医疗器械分别增加了 33% 和 27%。这些设备中超过 75% 用于放射学领域，可能是由于有大量的 predicate 设备，这可能为 510 (k) 许可提供更清晰的途径。此外，FDA 授权的设备使用预测性人工智能而非生成式人工智能（GenAI），后者还处于较新的阶段。

有关药物和生物制品临床开发的趋势，请参阅趋势（A）（1）；有关药物和生物制品临床开发用例的信息，请参阅本章后面的“医疗产品生命周期各步骤的示例用例和风险表”；有关药物和生物制品发现中人工智能的趋势和用例的信息，请参阅“医学研究与发现”章节，这可能是最普遍和成熟的应用领域。

2.5 医疗产品及其开发中人工智能的潜在用例和风险

以下对本章范围内的医疗产品生命周期的部分环节进行了描述，类似于本计划其他章

节中概述的“价值链”，以帮助指导后续关于用例和风险的讨论。请注意，医疗产品生命周期的临床前步骤（例如基础研究、发现）在“医学研究与发现”章节中进行了讨论。

图 6：本章范围内的医疗产品生命周期步骤



上图展示了本章范围内从临床开发到上市后产品安全监测的总体医疗产品生命周期。药物和生物制品以及医疗器械的开发过程遵循相同的总体步骤，但在这些步骤中的过程有所不同，特别是在监管批准方面。下面对展览中所示的医疗产品生命周期的每个步骤进行解释：

1. 临床开发在药物和器械之间有所不同，总结如下：

- a. **药物和生物制品**：在进行药物或生物制品的临床试验之前，必须向 FDA 提交药物和生物制品的研究性新药（IND）申请。从高层次来看，药物开发涉及一系列针对人类受试者的临床研究，以评估候选技术的安全性和有效性，一般分为三个阶段：I 期测试安全性和剂量；II 期评估初步疗效和安全性；III 期进一步评估疗效和安全性。在某些情况下，如某些疫苗或药物，FDA 可能要求进行 IV 期试验或上市后安全性研究，以进一步评估已知或潜在的严重风险。
- b. **医疗器械**：器械开发项目通常不遵循与药物相同的阶段顺序。如果特定器械在获得 FDA 市场授权之前需要在临床试验中进行测试，可能需要申请研究性器械豁免（IDE），尽管许多基于软件的器械研究风险不高，仅在机构审查委员会（IRB）的监督下进行。

2. **监管审查对于药物和生物制品与医疗器械可能有所不同**：鉴于审查过程的复杂性，本计划不会尝试总结步骤，而是指出 FDA 在这两方面的资源：

- a. **药物和生物制品**：药物和生物制品的开发和批准过程的详细描述可在脚注中查阅。
- b. **医疗器械**：医疗器械的市场授权过程的详细描述可在脚注中查阅。

3. **制造和供应链是指采购必要材料、使用它们开发医疗产品并在产品获得市场授权后向下游客户分销的运营过程**：制造商必须遵守适用的监管要求，包括 FDA 的质量体系法规 / 医疗器械现行良好生产规范（CGMP）以及药物和生物制品的 CGMP 法规，以确保医疗产品在生产过程中不被掺假。

4. **市场准入、商业和其他运营涉及开发和分发材料，解释产品在各种护理情况下对潜在提供者、支付者或其他利益相关者的相关性和影响**：这些包括物流、销售、定价、

财务、健康、经济学、结果研究和其他支持利益相关者的活动。FDA 监管医疗产品的营销，包括但不限于防止医疗产品的虚假或误导性标签。

5. 上市后安全和有效性监测包括在临床环境中使用医疗产品、持续监测其安全性并识别和缓解问题以确保患者持续安全：医疗器械上市后监测的要求包括报告器械故障、严重伤害或死亡，以及检查器械生产或分销的场所。对于药物，FDA 通过 FAERS 和哨点计划仔细监测其性能。此外，疫苗尤其通过各种监测系统（如疫苗不良事件报告系统、FDA BEST（生物制品有效性和安全性）计划和 CDC 的疫苗安全数据链）进行密切监测。

人工智能的应用在推动医疗产品创新和整个医疗产品生命周期方面具有巨大潜力，使患者受益，但在实施过程中应谨慎注意风险缓解。虽然与药物、生物制品和器械相关的人工智能风险有所不同，但随着技术的快速发展，一些重要的高层次主题浮现出来。在临床开发中，人工智能可能会延续其训练或调整数据中固有的偏差。在制造和供应链环节，当使用人工智能跟踪和管理供应链进行制造时，可能会因人工智能对供应需求的预测不准确而产生风险，导致生产不足。生产不足可能导致短缺，使人们无法获得对其护理至关重要的医疗产品。鉴于这些主题和以下描述的其他风险，HHS 已经在努力防范这些风险，并将继续探索潜在行动，以鼓励在该领域安全、创新地使用人工智能。

2.5.1 本章范围内医疗产品生命周期各步骤的示例用例和风险表

人工智能正在医疗产品生命周期中得到应用。在下表中，HHS 重点列出了本章范围内各步骤的潜在益处、用途和风险的详尽清单。各方应考虑适用的法定和监管要求，并在适当时咨询相关监管机构。请注意，以下用例突出了该领域各种利益相关者使用人工智能的现有或潜在方式。有关 HHS 及其部门如何使用人工智能的详细信息，请参考 HHS 人工智能用例清单 2024。

• 功能组件 1：临床开发

包括对研究性医疗产品进行人体参与者研究以评估其安全性和有效性

请注意，“医学研究与发现”章节讨论了基础研究和临床前开发，包括与人工智能在靶点识别、先导化合物和命中物生成及优化方面的用例和风险的讨论。

潜在用例（非详尽）	潜在风险（非详尽）
可帮助提高试验人群代表性的预测和分析模型 例如，为最大程度实现招募目标进行试验地点选择有助于识别有代表性患者的临床研究地点以助力达成招募目标 例如，进行候选者选择以确保试验人群具有代表性利用高级分析来确定代表若产品获批将会使用该产品的人群的队列	可能误导研究方向 例如，临床试验中的“假阳性”或“假阴性” 在辅助临床试验研究人员的技术中，人工智能可能会识别出并非真实的安全事件，或者未能识别出严重的安全事件。如果研究人

能够加速临床试验时间线的生成式、表征式和预测性模型 例如, 通过降低返工可能性来提高成功概率的临床试验设计策略 利用能够模拟潜在试验设计并推荐成功概率最高的子集的生成式和分析式模型 例如, 用于更快的计算机模拟实验的数字孪生 通过对产品进行虚拟模拟测试, 对物体、系统或候选对象进行虚拟表征可以加速研究	员过度依赖人工智能的特征描述, 或者在监督人工智能时出现人为错误, 这可能会导致错误分类, 并影响在分析数据时得出结论的能力。 可能存在偏差 例如, 使用医疗产品的人群缺乏代表性
例如, 终点评估和生物标志物识别 将人工智能作为临床结果评估的一部分或用于识别可能作为临床试验终点的生物标志物 256 例如, 图像、视频和语音分析以加速分析并可能提高其质量 利用人工智能 (通常是深度学习) 对成像数据、视频或语音进行分析, 有助于更快且可能更精确的分析 257 例如, 患者风险分层和剂量优化以提高试验参与者的安全性 根据“患者基线信息”预测特定严重不良事件的剂量和患者风险, 并随后使用此预测“帮助确定每个患者的住院或门诊监测需求”	虽然人工智能可以通过识别参与者、设计试验、分析结果等方式推进医疗产品开发, 但它可能没有在代表最终可能在临床上使用该医疗产品的人群的数据上进行训练。如果人工智能模型不是基于有代表性的数据进行训练, 这可能会导致研究结果仅与小部分人群相关, 并可能错失解决健康差距的机会。

- 功能组件 2: 监管审查
 - 向 FDA 提交文件

潜在用例 (非详尽)	潜在风险 (非详尽)
------------	------------

利用生成式模型加速医学写作的开发并提高其质量 例如，自动撰写临床研究报告（CSRs）以减少研究人员起草结果的时间 利用自然语言处理（NLP）和机器学习算法合成可在经人工适当确认后纳入向 FDA 提交的监管文件的结果 例如，生成所有可提交用于监管批准的文件的医学内容，以提高文件开发速度并在经人工适当确认后可能提高其质量 从现有材料生成研究或其他医学文档的初稿，以提高文档开发速度，并在经人工适当确认后有可能提升其质量。	可能存在不准确之处，降低批准机会 例如，患者或候选者记录与医疗保健专业人员（HCP）或研究人员笔记的错误合成 利用人工智能合成或生成与患者记录相关的内容（有时涉及人工编写的笔记）可能会导致输出不适合当前情况，因为数据质量差可能会导致输出不佳。使用此类工具可能需要对其分析的数据类型及其输出进行仔细监督。 可能引入安全风险 例如，从监管提交的研究结果中生成的见解并非基于数据 某些人工智能（例如大型语言模型）生成的内容可能是推断出来的而非基于事实，导致监管提交包含不准确信息。如果未被发现，此类不准确信息可能会导致不安全和无效的医疗产品获得市场授权
---	---

• 功能组件 3：制造和供应链

- 与产品采购、开发和向下游客户分销相关的运营

潜在用例（非详尽）	潜在风险（非详尽）
预测性和监测工具，可提前识别问题或低效情况 例如，监测制造运营以进行实时分析并提供行动建议以提高运营效率 接收药物生产过程的实时数据以提高生产率、纠正低效情况、控制质量和预测产量	可能扰乱关键医疗产品的供应 例如，人工智能赋能的供应链和制造运营监测可能会扰乱关键药物、生物制品和器械的运营 如果没有专家人工监督进行适当实施和管理，使用人工智能跟踪和管理原材料供应链可能会导致供应需求的人工智能预测不准确，从而导致生产不足。生产不

潜在用例（非详尽）	潜在风险（非详尽）
	足可能导致短缺，使人们无法获得对其护理至关重要的医疗产品
优化算法，有助于确保满足患者需求并减少短缺或产品浪费的可能性 例如，最大化现有物理和运营基础设施的生产输出 通过自动库存跟踪预测运营、人员和机器的性能以降低缺货和供应延迟风险	可能存在偏差 例如，医疗产品供应的不公平分配 利用人工智能规划药物和医疗器械需求的需求、物流和生产，如果人工智能分析中使用的数据不能充分代表相应产品所服务的患者人群，可能会导致分配不均。这可能会加剧现有的健康不平等并强化偏差，如果受影响的人群获得其健康所需的药物、生物制品和器械的机会减少

• 功能组件 4：市场准入、商业和其他运营

- 与潜在医疗服务提供者和支付者联系，解释医疗产品的相关性和影响（包括定价、财务、物流和支持活动）

潜在用例（非详尽）	潜在风险（非详尽）
分析性和生成式工具，可简化和加强市场进入活动 例如，为患者和 HCP 代表提供辅助工具以减少知识差距 利用在所有产品细节上进行训练的生成式人工智能快速回答患者和 HCP 代表可能不熟悉的主题的问题 例如，识别营销材料中的不准确信息 利用高级分析搜索互联网和其他推广医疗产品的资源，与 FDA 批准的标签进行比较并标记与营销相关的潜在监管问题	可能存在偏差 例如，创建不成比例地针对特定人群的营销战略和内容 如果分析工具未在代表性数据上进行训练，它们可能会限制某些可能已经服务不足的人群获得产品的机会

潜在用例（非详尽）	潜在风险（非详尽）
反馈和沟通工具，便于回答患者和医疗服务提供者关于医疗产品的问题并收集他们的意见 例如，HCP 参与和体验 自动回答 HCP 问题并提供动态反馈 例如，患者参与和体验 简化与患者的沟通并自动进行不需要人工解释的后续互动	可能会产生进一步加剧健康不平等的通信 例如，根据语音或写作模式生成的通信可能会进一步加剧健康不平等 利用生成式人工智能回答 HCP 和患者的问题或反馈，如果未在基于不同识字水平、方言、语言等的适当数据上进行训练，可能会导致有偏差或不准确的回答，从而延续现有的不平等

- 功能组件 5：上市后安全和有效性监测

- 在现实环境中监督和使用医疗产品以提供护理并持续监测产品安全

潜在用例（非详尽）	潜在风险（非详尽）
分析工具，可立即识别和报告医疗产品的疗效、安全和合规情况 例如，对医疗产品使用进行实时安全监测 分析临床数据以识别潜在的不良反应或其他安全信号，可能使快速响应以保护患者安全成为可能。 例如，对全国不良事件报告进行自动化分析和模式识别 对不良事件报告数据的高级分析模型用于识别临床环境中使用的医疗产品的潜在安全问题。 例如，简化药物警戒报告 根据医疗产品使用收集的数据对事件进行分类、自动生成反馈见解并识别新出现的问题 例如，持续合规监测	可能降低护理质量 在作为临床医生辅助工具的设备（例如筛查工具、人工智能辅助手术工具）中，人工智能在上市后监测和监控中可能会检测到与医疗产品安全无关的异常、副作用或不良反应，或者未能识别出合法的异常、副作用或不良反应。同样，利用人工智能进行的药物警戒分析可能会识别出“假阳性”或“假阴性”。虽然 HCP 和安全监测机构可以作为人工监督环节，但存在过度依赖人工智能或在解释人工智能时出现人为错误的风险，这可能会导致错误或不准确的安全报告

自动化合规审计并确保遵循标准操作程序 (SOPs)	
---------------------------	--

在医疗产品生命周期的每个阶段都有机会开发和使用人工智能来改善结果。各方都有责任在创新的同时监测和减轻风险。HHS 将使用以下行动计划，安全、负责、公平且有效地促进人工智能的使用。

2.6 行动计划

鉴于医疗产品及其开发中不断变化的人工智能格局，HHS 已采取了多项措施，包括提供监管清晰度、建立公私合作伙伴关系以及促进相应人工智能技术的公平性，以推动负责任的人工智能发展。以下行动计划遵循支持 HHS 人工智能战略的四个目标：1. 促进健康人工智能创新和使用；2. 推动可信人工智能开发和道德及负责任使用；3. 使人工智能技术和资源民主化；4. 培养人工智能赋能的劳动力和组织文化。对于每个目标，行动计划提供了背景、HHS 和其他相关联邦机构迄今为止的行动概述，以及 HHS 将采取的近期和长期优先事项。HHS 认识到，随着技术的发展，该行动计划需要不断修订，并致力于提供结构和灵活性，以确保产生长期影响。

2.6.1 促进健康人工智能创新和使用

为了抓住人工智能改变患者护理可及性和结果的机会，HHS 在推动医疗产品及整个医疗产品生命周期的创新和使用方面发挥着积极作用。HHS 有机会通过以下行动主题来增加该领域的人工智能应用：

1. 明确医疗产品的监管监督
2. 提供支付模式的清晰度
3. 促进公私合作伙伴关系和政府间合作，以快速开发和共享知识

以下，HHS 将更详细地讨论每个领域的背景、迄今为止的相应行动以及推进医疗产品人工智能创新和使用的计划

1. 明确医疗产品的监管监督

背景：

在医疗产品生命周期中可使用的人工智能的发展有了很大增长。特别是对于设备，新技术的强大功能和可用性的快速增长刺激了利用人工智能的健康信息技术应用的发展，这些应用超出了医疗器械法规的范围。《21 世纪治愈法案》(Cures Act) 特别从《联邦食品、药品和化妆品法案》(FD&C Act) 中删除了“设备”软件功能的定义，这些功能旨在：

- 为医疗保健机构提供行政支持

- 维持或鼓励与疾病或状况的诊断、治疗、缓解、预防无关的健康生活方式
- 作为电子病历
- 传输、存储、转换格式或显示测试或其他设备数据、结果或发现，但不旨在解释或分析它们
- 某些临床决策支持（CDS）软件
不被视为设备的 CDS 软件类型（“非设备 CDS”），例如支持或向 HCP 提供建议的应用程序，并且：
- 不获取、处理或分析医学图像、信号或模式
- 不显示、分析或打印超出医疗专业人员之间通常交流的医疗信息
- 不提供特定输出或指令
- 不要求医疗专业人员主要依赖建议，而是通过提供建议的基础来为决策提供信息

ASTP 的 HTI - 1 最终规则解决了某些经认证的电子健康记录（EHR）系统中人工智能的可用性问题，截至 2021 年，全国 96% 的医院和 78% 的医生办公室已使用这些系统。HTI - 1 最终规则本身并不创建审批流程，但它确实制定了政策，要求某些经认证的健康 IT 产品对其提供的技术保持透明。从 2025 年 1 月 1 日起，最终规则中确定的法规要求，对于作为健康 IT 产品一部分提供的、经认证符合 45 CFR 170.315 (b)(11) 的任何决策支持干预技术（包括基于人工智能的决策支持干预），必须提供特定的“源属性”信息。这些要求适用于基于人工智能的技术，无论设备定义、用例（例如临床与行政）或风险类别如何。随着健康 IT（例如 EHRs）中人工智能的持续增长，需要更明确的监管边界和适用性，以最大程度地减少可能阻碍创新的业务不确定性。虽然这个行动主题可能与设备比与药物和生物制品更相关，但人工智能的广泛可用性正在刺激所有医疗产品的增长。随着开发者做出投资和产品路线图决策，越来越需要进一步明确决定监管途径的定义，这些途径可能会影响设备、药物和生物制品开发的成本和时间。

HHS 迄今为止的行动（非详尽）：

- FDA 的《人工智能赋能设备软件功能：生命周期管理和市场提交建议指南》提供了关于包含人工智能赋能设备软件功能的设备的市场提交内容的建议，包括支持 FDA 评估安全性和有效性的文档和信息。这些建议反映了对设备整个产品生命周期（TPLC）中风险管理的综合方法。为了支持为 FDA 评估设备而开发适当的文档，该草案指南还提出了制造商在整个 TPLC 中可能希望考虑使用的人工智能赋能设备的设计、开发和实施建议。
- FDA 的《考虑使用人工智能支持药物和生物制品监管决策的指南》为申办者和其他相关方提供了关于使用人工智能生成旨在支持药物安全性、有效性或质量监管决策的信息或数据的建议。具体而言，该指南提供了一个基于风险的可信度评估框架，可用于建立和评估特定使用背景（COU）下人工智能模型的可信度。
- FDA 的《人工智能赋能设备软件功能预定变更控制计划的市场提交建议指南》为人工智能赋能设备量身定制了预定变更控制计划（PCCPs）的建议，并旨在通过对人工智能赋能设备的修改支持迭代改进，同时确保安全性和有效性。

- FDA 的《临床决策支持软件行业和 FDA 工作人员指南》澄清了《21 世纪治愈法案》中排除某些 CDS 软件受 FDA 设备管辖权的立法。这有助于阐明医疗保健技术中某些不受监管的人工智能使用的复杂性。
- FDA 的《人工智能与医疗产品：CBER、CDER、CDRH 和 OCP 如何合作》文件详细说明了生物制品评估和研究中心（CBER）、药物评估和研究中心（CDER）、设备和放射健康中心（CDRH）以及组合产品办公室如何合作，以确定促进合作、制定法规、推广最佳实践和支持相应研究工作的步骤。
- FDA 的《数字健康和人工智能术语表 - 教育资源》是一个公开可用的资源，定义了数字健康、人工智能和机器学习中的常用术语，以提供内部和外部的一致性和教育。
- AHRQ 的临床决策支持创新协作一直在推进以患者为中心的临床决策支持（PC CDS），包括探索人工智能对 PCCDS 的影响并开展试点项目。

HHS 近期优先事项：

- 继续发布关于医疗产品开发和医疗产品中人工智能使用的指南、支持材料（例如常见问题解答）和 / 或讨论文件，以提供进一步的建议。
- 考虑新的资源机会来研究人工智能和 CDS，包括更好地理解在 CDS 软件中使用临床数据的益处和风险的方法。

2. 明确支付模式

背景：

在各个临床学科（例如放射学和病理学）中，一些设备结合了已证明有效的人工智能；然而，由于许多设备没有既定的支付方式，其全部使用潜力尚未实现。医疗保健支付和覆盖政策会影响人工智能使用的经济基础。虽然一些医疗产品在有市场的情况下可能具有明显的效率或生产力投资回报效益，但在美国复杂的医疗支付方式中，患者利益和财务激励之间可能存在脱节。购买者和支付者需要与覆盖决策相关的患者人群的结果和 / 或终点证据，以及与支付者相关且有利于商业化和患者可及性的使用指征。如果在临床环境中没有明确的使用途径，医疗设备开发者可能缺乏继续创新此类产品的动力。

一般来说，要考虑医疗保险覆盖一项物品或服务，该物品或服务必须至少属于《社会保障法》（the Act）中规定的一个福利类别，该物品或服务不得被该法案明确排除，并且该物品或服务必须“对疾病或伤害的诊断或治疗是合理和必要的”。CMS 可能会发布国家覆盖决定（NCD）来描述医疗保险对特定物品或服务的全国覆盖条件。如果没有 NCD，物品和服务将根据医疗保险行政承办商（MACs）的决定或通过地方覆盖决定逐案进行覆盖。截至 2024 年 5 月，CMS 已通过现行程序术语（CPT®）和新技术附加支付（NTAP）在医疗保险住院预付费系统（IPPS）下为至少八种人工智能赋能设备建立了支付途径，不到 FDA 授权的基于人工智能产品的 5%。CMS 还在 2022 年通过单独的软件即服务附加代码为医院门诊部门建立了支付途径。随着时间的推移，

基于价值的采购支付模式的增长可能会为医疗保健领域的人工智能投资提供更多内在的财务激励，但此类计划的增长并不迅速。进一步澄清现有途径可以促使更多人工智能赋能设备获得既定支付。

HHS 迄今为止的行动（非详尽）：

- CMS 的 NTAP 为医疗保险住院预付费系统（IPPS）下的某些新设备（包括利用人工智能的设备）提供附加支付，一些例子可以追溯到 2020 年（例如 Viz.ai 的 ContactCT，一种用于大血管闭塞的人工智能驱动的分诊软件）。从那时起，其他人工智能软件开发者（例如 RapidAI、Aldoc、Avicenna）也获得了 NTAP 地位。
- CMS 的新兴技术过渡性覆盖（TCET）途径帮助医疗保险受益人获得最新的医疗进展，使医生和其他临床医生能够为他们的患者提供最佳护理，并使创造创新技术的制造商受益。
- CMS 的医疗保险药品和技术监察员自 2017 年末以来一直存在。该监察员接收并协助处理制药、生物技术、医疗设备、诊断产品制造商和其他利益相关者关于医疗保险覆盖、编码和 / 或支付的询问和投诉，或寻求医疗保险覆盖的产品。

HHS 近期优先事项：

- 召集 HHS 部门（例如 CMS、NIH、FDA 和 ASTP），就人工智能赋能医疗设备的标准化、面向未来的支付途径的益处、风险和潜在定义达成一致。
- 扩大早期支付者反馈计划，缩短与商业和政府保险公司的支付和覆盖决定时间。
- 发布关于医疗保健支付者、提供者和其他利益相关者建立人工智能赋能设备支付途径的指南。

HHS 长期优先事项：

- 为公共部门的人工智能赋能医疗设备开发明确的支付途径，以潜在地刺激私营部门的类似活动。
- 随着医疗技术的转变，迭代地重新评估人工智能赋能医疗设备的指南和支付途径，以继续促进使用并减轻风险。

3. 促进公私合作伙伴关系和政府间合作，以快速开发和共享知识

背景：

世界各地的监管机构正在采取不同的方法发布关于医疗产品中人工智能的指南。旨在为全球患者提供服务的医疗产品开发者、制造商和分销商可以通过遵循

合作性的标准和保障措施更有效地进行创新。确保医疗产品安全和有效的监管流程对于保护美国公众至关重要，FDA 的医疗产品中心打算继续管理相关项目，以加速创新并为人工智能的使用提供监管指南。此外，通过继续并加强与私营部门的直接互动，HHS 能够以一种促进医疗产品及整个医疗产品生命周期中人工智能进一步发展的方式共享知识。

HHS 迄今为止的行动：

- FDA 通过与其他政府、学术、科学、患者和私营部门组织的合作参与公私合作伙伴关系 (PPPs)，推动了医疗产品开发、评估和制造方面的科学与创新。这些持续的努力鼓励开发包括人工智能在内的新工具，以促进医疗产品生命周期的创新。一些可能聚焦人工智能的公私合作伙伴关系示例如下：
 - BioFabUSA 致力于将创新的细胞和组织培养与生物制造、自动化、机器人技术和分析技术的进步相结合，创建具有颠覆性的研发工具和符合 FDA 要求的批量制造流程。
 - 国家生物制药创新研究所 (NIIMBL) 促进创新制造技术和劳动力发展项目，以提高生命科学行业的效率和影响力。
 - 关键路径研究所 (C - Path) 是一个非营利组织，致力于通过促进私营部门行业高管、科学家、学术研究人员、监管机构和患者团体之间的合作来改进和简化药物开发。
 - 临床试验转型倡议 (CTTI) 汇聚了代表学术界、临床研究人员、政府和监管机构、私营部门行业、机构审查委员会 (IRB)、患者倡导团体等的组织和个人，共同开发基于证据的临床研究挑战解决方案。
- NIH 的通过道德、多模式人工智能 (AI) 促进健康研究倡议资助开发以道德为重点且数据驱动的多模式人工智能方法，以更紧密地解释和预测复杂的生物和行为系统，并对复杂的健康系统进行建模，从而增强我们对健康的理解以及检测和治疗人类疾病的能力。
- FDA 的人工智能项目 —— 基于人工智能的医疗设备研究依靠 CDRH 进行监管科学研究，以确保患者能够获得使用人工智能的安全有效的医疗设备。具体重点领域包括增强模型训练的方法、最小化偏差以及开发跟踪上市后安全的方法。
- FDA、NIH 和 NSF 启动了数字孪生作为生物医学技术创新催化剂 (FDT - BioTech) 计划，通过合成数据催化生物医学创新，通过提供传统参与者招募可能难以获得的对照数据来促进临床试验。
- NIH 及其各个研究所、中心和办公室正在资助研究，以将人工智能应用于许多疾病背景中，包括利用可穿戴技术帮助监测和筛查认知障碍、通过视网膜成像检测神经系统疾病以及识别潜在的药物滥用障碍患者。
- NIH 的国家癌症研究所 (NCI) 与能源部 (DOE) 的合作作为癌症登月计划的一部分，加速了精准肿瘤学和科学计算的进步，包括人工智能的使用。

HHS 近期优先事项：

- 利用并继续拓展现有的围绕医疗产品及整个医疗产品生命周期中人工智能使用的倡议。

- 探索公私合作伙伴关系的方法，以推进医疗产品中人工智能的创新、商业化和风险缓解方法，并促进该领域安全、负责、公平、保护隐私和可信的人工智能发展，如第 11410 号行政命令所阐述。
- 评估继续扩大总产品生命周期咨询试点（TAP）和早期支付者反馈计划（EPFP）的方法，以加速识别人工智能的创新、使用和商业障碍，特别是对于那些不熟悉设备市场授权流程和支付者覆盖决策的开发者。
- 与针对投资不足的治疗领域的战略投资进行协调。

HHS 长期优先事项：

- 继续监测和评估趋势及新出现的问题，以发现潜在的知识差距和机会，从而及时调整，为医疗产品生命周期中人工智能的使用提供清晰指导。
- 继续与全球合作伙伴密切合作，促进在标准、指南和最佳实践方面的国际合作，以鼓励在医疗产品领域使用和评估人工智能的合作。
- 探索资源开发教育倡议，以支持监管机构、医疗保健专业人员、患者、研究人员和私营部门行业在医疗产品及其开发中安全和负责地使用人工智能。
- 探索资源支持监管科学努力，开发额外的评估人工智能算法的方法，识别和缓解偏差，并确保其对不断变化的临床输入和条件的稳健性和适应性。

2.6.2 促进可信人工智能开发和道德及负责任使用

HHS 将通过以下方式促进医疗产品中或整个医疗产品生命周期中人工智能的可信、道德和负责任使用：

1. 完善监管框架以应对医疗设备中的自适应人工智能技术
2. 促进人工智能部署中的公平性以支持安全和负责任使用
3. 解决当前设备监管机构管辖范围外的人工智能赋能软件问题
4. 培育健康人工智能的私人或公共质量保证机制

以下，HHS 将更详细地讨论每个领域的背景、迄今为止的相应行动以及确保医疗产品中人工智能使用可信和安全的前瞻性计划。

1. 完善监管框架以应对医疗设备中的自适应人工智能

背景：

FDA 的传统医疗器械监管模式可能并非为自适应人工智能技术而设计，这些技术能够实时持续改变和优化设备性能以改善患者医疗保健。当前的监管方法是在市场授权时监测设备的配置性能和安全，可能无法应对诸如人工智能等自适应技术，这些技术可能与最初授权时的情况有很大偏差。大多数 FDA 授权的医疗设备通过 510 (k) 途径获得授权，该途径基于证明与合法上市的“predicate”设备实质等效。随着此类技术的复杂性增加，对这些产品的安全性和有效性进行更具体和明确的上市前演示可能

有助于考虑自适应人工智能和其他技术。这些工具的高度迭代、自主和自适应性质可能受益于总产品生命周期 (TPLC) 方法，这是一种监管方法，有助于快速的产品改进周期，并允许这些设备在持续提供有效保障的同时进行改进。通过适当定制的监管监督，人工智能可以提供安全有效的功能，提高患者护理质量。

HHS 迄今为止的行动（非详尽）：

- FDA 2021 年的《基于人工智能和机器学习的软件作为医疗设备 (SaMD) 行动计划》概述了一种多管齐下的方法来推进该机构对这些技术的监督。FDA 已经：
 - 发布了关于人工智能赋能设备软件功能预定变更控制计划的市场提交建议草案指南。
 - 与加拿大卫生部和英国药品和医疗产品监管局 (MHRA) 的合作伙伴共同发布了《医疗器械开发良好机器学习实践指导原则》。
 - 主办了关于人工智能赋能医疗设备透明度的公共研讨会。
 - 发布了《聚焦：数字健康监管科学机会》。该聚焦突出了包括人工智能和机器学习在内的常见数字健康兴趣领域等主题，并向所有人展示了数字健康领域当前的监管科学兴趣领域以供考虑。
- ARPA - H 的人工智能持续修改和可用性性能与可靠性评估 (PRECISE - AI) 计划资助研究开发技术，以检测临床护理环境中使用的人工智能赋能工具何时与基础训练数据不一致并自动纠正它们。

HHS 近期优先事项：

- 探索使用人工智能为监管决策生成信息的政策，包括定义感兴趣的问题、使用背景、模型风险和模型输出可信度的潜在方法。
- 探索在各种监管框架中应用“模型卡”方法来管理人工智能。
- 开发评估人工智能赋能医疗设备的安全性、有效性和 / 或性能的标准、指南和创新的基于科学的方法。
- 探索资源用于研究评估和监测医疗设备中的人工智能性能。
- 探索资源用于评估和使用强大的人工智能工具来模拟医疗设备中的漂移，作为对 ARPA - H PRECISE - AI 程序的潜在补充。
- 将人工智能纳入申办者的监管提交和 FDA 内部审查流程。

HHS 长期优先事项：

继续完善和发展在医疗产品生命周期中评估人工智能安全、有效、负责和道德使用的考虑因素（例如，人工智能提供足够的透明度并解决安全、有效性和网络安全问题）。

2. 促进人工智能部署中的公平性以支持安全和负责任使用

背景：

FDA 正在采取措施在医疗产品背景下推进健康公平。ASTP 对认证健康 IT 产品的要求确实包括健康公平组件；然而，ASTP 法规的范围仅限于认证健康 IT 或产品，包括认证健康 IT。随着医疗产品中及整个医疗产品生命周期中人工智能的使用不断增加，HHS 可以考虑采取方法来增强该领域的健康公平性。

HHS 迄今为止的行动：

- FDA 的《人工智能与医疗产品：CBER、CDER、CDRH 和 OCP 如何合作》文件讨论了 FDA 的医疗产品中心如何与开发者、患者团体、学术界、全球监管机构和其他利益相关者密切合作，培育以患者为中心的监管方法，强调合作和健康公平。该文件还描述了 FDA 对考虑医疗产品开发中使用人工智能相关健康不平等问题的项目的支持，以促进公平并确保数据代表性，利用持续的多样性、公平性和包容性努力。
- ASTP 的博客文章《通过设计拥抱健康公平》讨论了医疗保健 IT 公平性的多方面方法。它包括使用正确的数据、选择适当的工具以及确保系统之间的互操作性，以减少偏差并确保所有群体在健康技术中得到适当代表。
- AHRQ 的《数字医疗公平框架及实施实用指南》帮助组织在开发和使用数字医疗技术和解决方案时有意考虑公平性。该指南是数字医疗开发者、供应商、医疗系统、临床提供者和支付者的资源。它包括在数字医疗生命周期各阶段推进公平性的步骤和实际示例。

适用的联邦法律迄今为止：

- 《平价医疗法案》第 1557 节禁止在某些健康计划和活动中通过患者护理决策支持工具（包括人工智能）基于种族、肤色、国籍、性别、年龄和残疾进行歧视。（有关其他非详尽的联邦政策和法规，请参阅附录 B）

HHS 近期优先事项：

- 探索资源用于内部和外部项目，突出人工智能开发生命周期中可能引入偏差的不同点以及如何通过风险管理解决这些问题。
- 传播关于记录和确保用于训练和测试人工智能模型的数据适合使用并充分代表目标人群的最佳实践研究，以帮助增强促进安全和负责任人工智能使用的公平性考虑。
- 探索资源用于考虑医疗产品开发中使用人工智能相关健康不平等问题的项目，以促进公平并确保数据代表性，利用持续的多样性、公平性和包容性努力，以帮助确保医疗产品及其开发中人工智能的道德和可信使用。
- 探索资源用于利用人工智能进行临床试验，以解决未满足需求的领域或管道未满足需求的情况。

HHS 长期优先事项：

- 继续探索资源用于内部和外部项目，突出人工智能开发生命周期中可能引入偏差的不同点以及如何通过风险管理解决这些问题。

3. 解决当前设备监管机构管辖范围外的人工智能赋能软件问题

背景：

越来越多的健康 IT 中的人工智能工具可能不在 FDA 监管范围内，包括某些集成到 EHR 中的人工智能决策支持工具（例如预约未到预测）以及健康计划和保险发行商为使用管理和预先授权部署的人工智能算法。对非医疗设备的健康 IT 的监管权部分属于 ASTP/ONC。不符合 FDA 设备定义的工具可能不会接受监管审查、验证或测试。这是 HHS 将继续密切监测的一个领域。

HHS 迄今为止的行动：

- ASTP/ONC 的 HTI - 1 最终规则最终确定了政策，要求某些认证健康 IT（例如符合 45 CFR 170.315(b)(11) 认证标准的 EHR 健康 IT 产品）使用户能够访问有关人工智能（称为预测决策支持干预或 PDSIs）的设计、开发、培训和评估的信息，以帮助用户确定该工具是否适合其护理环境和患者人群。
- FDA 的《临床决策支持软件行业和 FDA 工作人员指南》澄清了《21 世纪治愈法案》中排除某些 CDS 软件受 FDA 设备管辖权的立法，这有助于阐明医疗保健技术中某些不作为设备监管的人工智能使用的复杂性。

HHS 近期优先事项：

- 评估确保对 FDA 监管范围外的人工智能进行适当监督的机制，并持续监测生态系统的进展。

探索以下方法：

- 为 FDA 管辖范围外的基于人工智能的技术提供“模型卡”信息。
- 加强对基于人工智能的模型的临床数据验证。
- 在监管途径中纳入健康公平考虑因素。
- 建立公私合作模型，对 FDA 管辖范围外的基于人工智能的技术进行严格的、基于标准的上市前和上市后质量保证。

HHS 长期优先事项：

- 随着该领域的快速发展，迭代地监测和重新评估 FDA 管辖范围外的医疗和健康技术中人工智能的监管监督机制。
- 探索收集关于 FDA 管辖范围外的医疗和健康技术中人工智能的反馈的机会，以监测此类技术对医疗保健的潜在影响。

4. 培育健康人工智能的私人或公共质量保证机制

背景：

尽管人工智能工具在医学中具有潜力，但在不同数据集上前瞻性地测试人工智能工具并在多个临床护理环境中部署人工智能以确保在改善健康结果方面的一致性、准确性和通用性的能力可能会受到此类数据集的可用性以及临床使用中不一致监测的限制。通过增加数据可用性和改进监测能力，可以改善对人工智能的测试，以识别潜在偏差、差异或人工智能模型性能的不一致，并优化人工智能模型以适应不同的医疗环境。缺乏旨在评估现实世界环境中性能以确保持续的患者和提供者安全的标准化质量保证（QA）协议，增加了在不同地点实施不一致和产生意外后果的风险。即使是获得临床使用监管批准的人工智能工具，在新的临床环境中部署时也可能表现不佳，这是由于泛化能力差或用于非授权预期用途。这些情况凸显了人工智能工具在医学中面临的挑战，这些挑战源于开发数据中的偏差（例如，训练、调整、开发者创建工具时使用的内部测试集）以及外部、以前未使用的测试集或患者病例特征的潜在分布变化。为了将人工智能工具安全有效地集成到临床工作流程中，制造商对开发过程的“透明度”以及实施 QA 计划可能是必要的。

HHS 迄今为止的行动：

- FDA 与退伍军人事务部于 2024 年 10 月宣布的合作将成为医疗相关人工智能工具的“机构间测试场”。据退伍军人事务部卫生副部长 Shereef Elnahal 称，该实验室“将成为联邦机构和私营部门的资产，使其能够在虚拟实验室环境中测试人工智能的应用，不仅确保它们有效且对退伍军人和患者安全，而且确保它们遵循可信人工智能原则”。

HHS 近期优先事项：

- 与公共和私人网络合作测试健康人工智能，提供共享资源和基础设施，鼓励安全和有效的开发、透明度、报告和持续监测健康人工智能。
- 考虑支持指南和教育工具，以帮助人工智能开发者在为医疗产品及整个医疗产品生命周期创建人工智能技术时朝着安全、安保和信任的方向努力。

HHS 长期优先事项：

- 探索资源开发监管科学方法，以评估部署在医疗环境中的人工智能模型的准确性和可靠性。

2.6.3 使人工智能技术和资源民主化

为了在减轻相关风险的同时有效捕捉医疗产品整个生命周期中人工智能的价值，技术的使用和创新可以受益于整个生态系统中不同参与者（例如医疗技术公司、学术界、非营利组织和公共部门实体）和利益相关者（例如来自不同人口背景）的公平获取。如果没有这种可及性，可能无法充分实现该领域人工智能的全部价值潜力或充分考虑风险。HHS 计划通过将公平原则融入医疗产品中人工智能的扩展，在以下关键行动主题中发挥关键作用来缓解这一问题：

1. 通过公众参与实现协作开发
2. 统一研究和医疗保健交付的标准和信息共享机制

以下, HHS 将更详细地讨论每个行动主题的背景、迄今为止的相应行动以及确保医疗产品中人工智能技术和资源公平获取的计划。

1. 通过公众参与实现协作开发

背景:

利益相关者之间的合作增加可以使医疗产品中的人工智能技术和最佳实践民主化, 并贯穿整个医疗产品生命周期。利益相关者(例如私营部门行业、STLTs、学术界和公众)之间缺乏合作以及在医疗产品生命周期中缺乏有意的公众参与可能会限制人工智能在医疗产品及其开发中广泛公平使用的潜力。

HHS 迄今为止的行动:

- NIH 的 AIM - AHEAD 计划旨在支持三方互利伙伴关系, 包括(1) 地方、州和部落认可的卫生部门;(2) 资源有限的高等教育机构;(3) 具有可访问数据库的数据科学导向组织, 共同开展与健康公平相关的人工智能研究。这些关键且可信的组织可以从增强其人工智能能力中受益, 以推进公共卫生, 从早期检测和监测、预测分析、疾病监测和监测到医疗资源分配和个性化干预。公共卫生部门专业人员、学术研究人员和数据科学 / 人工智能专家之间的伙伴关系可以进一步利用数据驱动的意见, 为更有效和高效的公共卫生策略做出贡献, 以改善社区健康结果。
- 能源部和 NIH 通过国家人工智能研究资源 (NAIRR) 安全试点的合作将“使涉及敏感数据的研究成为可能, 这些数据需要特殊处理和保护。NAIRR 安全试点将组装示例隐私 / 安全保护资源(例如数据保护区、安全计算资源和隐私保护工具) 并制定未来 NAIRR 安全的要求”。

HHS 近期优先事项:

- 制定在医疗产品生命周期的所有部分纳入公众声音的愿景和框架。
- 召集一个公私实践社区, 分享最佳实践并识别临床研究中人工智能使用的促进因素 / 障碍。
- 完善和发展更强大的 STLT 参与医疗产品战略, 在适当情况下确保各级政府之间共享人工智能最佳实践。

HHS 长期优先事项:

- 提供安全沙箱, 鼓励在医疗产品开发和使用中进行协作创新。
- 参与公私合作, 促进私营部门行业、提供者和公众之间的长期关系, 以便进行共同创造。

- 探索资源用于多机构合作机制，特别是那些可能资源不足但可以从知识或基础设施共享中受益的组织。

2. 统一研究和医疗保健交付的标准和信息共享机制

背景：

清晰的数据、元数据标准和信息共享途径可以使人工智能创新更容易获得。缺乏清晰的标准可能会使来自私营部门、学术界、非营利组织、政府和其他参与者的数据无法用于人工智能模型，从而阻碍医疗产品及整个医疗产品生命周期中人工智能的使用。对于资源获取较少的利益相关者来说，数据共享的障碍对创新的阻碍可能比对资源丰富、能够资助数据收集或清理活动的利益相关者更大。

HHS 迄今为止的行动（非详尽）：

- ARPA - H 与 FDA 的 CDRH 合作的成像数据伙伴关系旨在简化对负担得起的高质量医疗成像数据的访问。这些机构共同努力开发一个医疗成像数据市场，通过消除获取符合监管质量标准且能适当代表美国相关人群的数据的障碍，加速人工智能和机器学习创新。
- ARPA - H 的生物医学数据结构工具包寻求促进来自数千个来源的生物医学研究数据的连接，推进来自数千个不同研究实验室、临床护理中心和其他数据源的生物医学数据集的收集和可用性，并加速整个健康生态系统的技术创新。通过（1）降低以计算机可读形式进行高保真、及时数据收集的障碍，（2）为大规模多源数据分析做准备，（3）推进直观的数据探索，（4）在维护隐私和安全措施的同时改善利益相关者的访问，以及（5）确保生物医学数据结构工具在不同疾病类型中的通用性，ARPA - H 正在使数据访问民主化。这些数据必须是可查找、可访问、可互操作和可重用的。NIH 的通用存储库生态系统倡议 (GREI) 支持七个已建立的通用存储库，它们共同努力建立一致的元数据，开发数据共享和重用的用例，并培训和教育研究人员如何共享和重用数据，包括用于人工智能的开发和使用。
- NIH 的迈向生物医学和行为研究人工智能道德框架：数据和模型重用透明度研讨会重点强调了标准化合成数据、一般重用数据和多模式数据的安全可共享性的重要性，如果在人工智能工具中加以利用，这些数据可能会带来变革性的产品开发。

HHS 近期优先事项：

- 发布符合 HHS 数据战略的数据共享原则草案指南，包括构建数据和元数据的通用方法以及明确可发布和共享的数据类型。
- 提供安全沙箱，以促进数据共享和标准开发方面的合作。

- 开发开放行业标准、开源工具和基础设施用于注册，以利用人工智能支持设备上市前和上市后提交要求、跨标准数据映射和去识别，以开发适合人工智能的数据集和工具。
- 加速与标准开发组织的合作以及行业在标准方面的协作，以支持人工智能在整个生命周期中的开发和使用。
- 加速联邦资助的研究数据标准（语义、格式和传输）与 HHS 使用的电子健康记录（EHR）、医疗服务提供者和支付者标准（例如 USCDI、USCDI +、HL7、FHIR 和 CARIN）的对齐。

HHS 长期优先事项：

随着公众获取研究结果、数据管理和共享的格局发生变化，HHS 可能需要增强能力，以协助关键参与者完善这两方面的标准。

2.6.4 培养人工智能赋能的劳动力和组织文化

如果没有足够的人工智能人才供应来实现医疗产品及整个医疗产品生命周期的大规模创新，广泛的使用和有效应用可能无法实现。为此，HHS 计划通过关注以下关键行动主题，在内部和外部促进劳动力发展，以支持医疗产品生命周期中持续的负责任、安全的人工智能创新：

1. 改善医疗产品中人工智能治理和管理的培训
2. 培养和留住与医疗产品相关的人工智能人才

以下，HHS 将更详细地讨论这一目标的背景、迄今为止的相应行动以及培养医疗产品中人工智能赋能的劳动力和组织文化的计划。

1. 改善医疗产品中人工智能治理和管理的培训

背景：

大多数参与人工智能的人员将负责管理和使用此类技术，而不是开发它们。确保医疗产品生态系统（包括开发者、临床医生和患者）充分利用人工智能不仅需要关注技术本身，还需要关注其实施、工作流程集成和生命周期管理。为研究人员提供培训，使其能够负责任地管理和使用这些技术，对于利用人工智能推进医疗产品至关重要。

HHS 迄今为止的行动（非详尽）：

FDA 的博客文章《实现安全、有效人工智能赋能医疗保健的生命周期管理方法》概述了一种开发、验证和管理医疗设备中人工智能使用的持续治理的潜在方法，以维护其安全性和有效性。这种方法可以为 HHS 提供基础，以进一步发展关于医疗设备中及开发过程中人工智能治理和管理培训的最佳实践。

HHS 近期优先事项：

探索针对资源、培训和研讨会，将临床研究中人工智能技术的治理和管理纳入其中，包括临床试验设计和管理。

HHS 长期优先事项：

通过有针对性的内部培训或学徒计划，培养与医疗产品相关的内部数据科学、计算机科学和人工智能人才。

2. 培养和留住与医疗产品相关的人工智能人才**背景：**

为了利用人工智能的潜力，私营部门行业、政府、学术界、非营利组织和其他相关方可能需要一个强大的渠道来培养多样化的劳动力，能够开发和嵌入人工智能以增强医疗产品及其开发。来自所有背景的专业人员都需要基础知识来安全、负责和有效地开发和应用人工智能。因此，培养和留住与医疗产品相关的人工智能人才对于保持创新增长至关重要。

HHS 迄今为止的行动：

- FDA 的 STEM 外联、教育和参与计划旨在为潜在科学家提供教育机会，提高对 FDA 作为科学机构的认识，让学生接触到广泛的监管科学及其对我们生活的影响，激励未来的创新者追求构成 FDA 监管科学领域的广泛科学职业，并招募和雇佣科学家。虽然该计划一般面向 FDA，但它增强了整体人才生态系统，并可以探索与人工智能相关的额外重点。
- NIH 的 Bridge2AI 计划基于道德原则、相关标准和工具以及技能和劳动力发展创建旗舰数据集，以应对需要人工智能分析的生物医学和行为研究中的重大挑战。
- FDA 的科学实习和研究金为本科生和研究生提供机会，探索与研究、监管科学和其他 STEM 领域相关的职业，这些领域培养了未来 FDA 和其他技术人才。虽然该计划一般面向 FDA，但它有助于促进对医疗产品中人工智能相关额外重点的探索。
- HHS 将人工智能整合到企业活动中（见内部运营章节）并发布了所有用例的公共跟踪器。截至 2023 年，HHS 及其部门有 164 个人工智能用例，包括数据去重、不良事件检测、安全监测、信号检测管理、数据可视化和文本分析。例如，FDA 正在探索在各个领域使用人工智能，包括在 FAERS 中对非公共不良事件数据进行去重，以及使用术语识别和新型合成阿片类药物检测与评估分析工具识别阿片类药物相关的新术语，该工具使用公开可用的社交媒体和法医化学数据在社交媒体文本中识别药物产品的新指代。

HHS 近期优先事项：

- 扩大实习和学徒计划，纳入与医疗产品及其开发相关的人工智能特定角色。
- 探索为现有外联、教育和参与计划提供额外资源，纳入与医疗产品及其开发相关的人工智能特定内容，特别是那些相关的内容。

- 评估扩大 NIH 的 AIM - AHEAD 计划，以包括临床研究中人工智能专业知识的招募和培训。

2.7 结论

人工智能可以成为医疗设备、医疗设备的一部分、增强临床试验的设计和实施、简化制造和供应链，并加强医疗产品的上市后监测和监控，最终改善患者护理和创新医疗产品的可及性。然而，人工智能的快速发展也带来了应加以解决的挑战。HHS 的平衡方法旨在促进人工智能创新，同时维持强大的监管框架，确保医疗产品保持安全、有效和高质量。

3 医疗保健交付

3.1 引言和背景

美国的医疗保健交付 —— 在此定义为医疗保健的融资、直接提供患者护理、相关行政服务和研究 —— 是一个庞大而高度复杂的系统。2022 年，美国的国家卫生支出（包括公共卫生）约为 4.5 万亿美元，占美国经济的 17%，并为约 9% 的全国劳动力提供了就业机会。在美国，医疗保健由持牌提供者提供，主要由支付者资助（例如，2022 年，92% 的美国患者有医疗保险）。除了患者、提供者和支付者之外，一系列利益相关者参与了医疗保健交付生态系统，包括提供支持护理的资源和技术实体。许多 HHS 实体，包括 CMS、HRSA、SAMHSA、IHS、AHRQ 等，直接参与促进医疗保健交付或为交付合作伙伴提供指南、支付和资金、培训及其他运营支持。

在医疗保健交付中，人工智能尤其有潜力增强从护理交付到医疗保健财务再到研究（例如卫生服务和行为健康）等广泛的活动。HHS 致力于最大限度地发挥人工智能对医疗保健交付系统中利益相关者的潜在益处 —— 为此，人工智能干预措施必须以患者为中心，将透明度、安全、公平和安全置于实施考虑的前沿。同样至关重要的是，通过确保新技术得到负责任的测试、部署和监测来保护美国人的安全。在接下来的章节中，HHS 概述了其针对医疗保健交付的四个目标和行动：（1）促进健康人工智能创新和使用；（2）促进可信人工智能开发和道德及负责任使用；（3）使人工智能技术和资源民主化；（4）培养人工智能赋能的劳动力和组织文化。

3.1.1 行动计划摘要

本章后面，HHS 阐述了为推进其在该部门负责任使用人工智能的四个目标而提议的行动。以下是每个目标内行动主题的摘要。有关提议行动的详细信息，请参阅第 3.6 节行动计划。

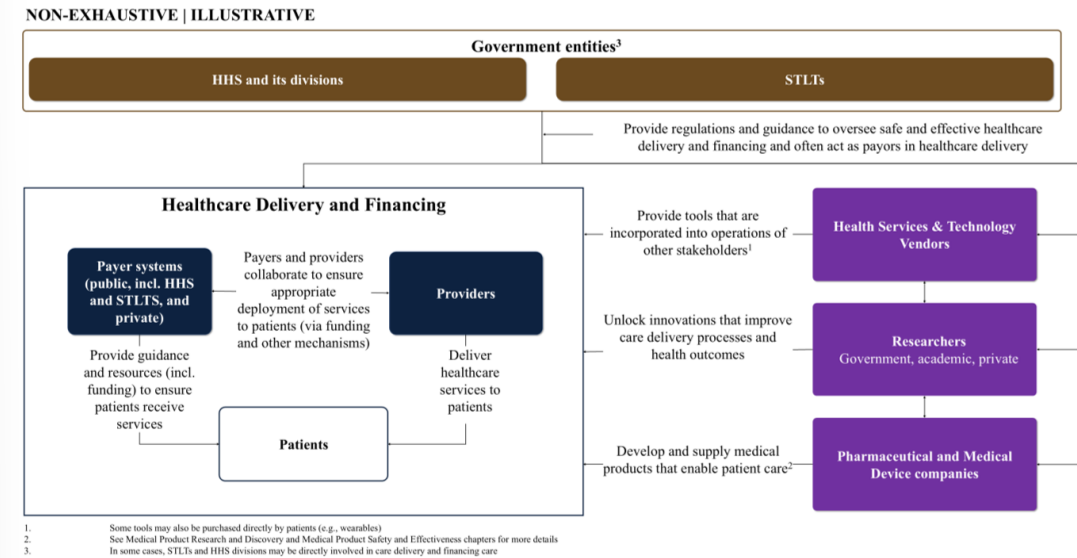
关键目标及支持行动	提议行动主题（非详尽，详见 3.6 行动计划）
1. 促进健康人工智能创新和使用	支持收集人工智能干预措施的有效性、安全性和风险缓解证据以及在医疗保健交付环境中实施的最佳实践 提供关于监督、医疗责任以及隐私和安全保护的指南和资源，以增强组织开发人工智能的信心 确保人工智能开发者和潜在部署者清楚覆盖和支付确定流程，以鼓励人工智能开发
2. 促进可信人工智能开发和道德及负责任使用	加强执法并澄清与现有要求相关的指南 提供与组织治理相关的指南和支持 促进外部评估、监测和透明度报告 增强基础设施以确保患者安全
3. 使人工智能技术和资源民主化	通过为服务不足人群提供服务的交付组织提供技术支持和合作，促进公平获取 为医疗保健交付组织提供支持，以解决核心基础设施和部署挑战（即技术、基础设施和数据基础设施），提高人工智能准备程度
4. 培养人工智能赋能的劳动力和组织文化	为医疗保健交付专业人员提供培训、资源和研究机会，以支持他们在各自医疗系统组织中的人工智能素养和专业性知识

3.2 参与医疗保健交付人工智能价值链的利益相关者

医疗保健交付是一组高度复杂的活动，涵盖通过公共或私人医疗保险为医疗保健融资以及通过私立和公立医院及门诊设施提供医疗服务。雇主和个人通过各种实体购买医疗保险。医疗保健由数千家医院和数百万临床医生及其他医疗保健专业人员提供，他们提供各种服务，并受到联邦和州、地方及部落政府（STLT）实体的监管。

图 7 展示了利益相关者之间的示例流程以及参与医疗保健交付的利益相关者的项目符号列表。请注意，该图和列表都没有涵盖所有利益相关者角色和互动。有关监管指导和权力的更多详细信息，请参阅其他 HHS 文件。角色可能因医疗保健交付系统或活动而异。

Exhibit 7: Healthcare Delivery Stakeholder Engagement Map



卫生与公众服务部（HHS）各部门及在医疗服务提供中的示例职责（非详尽列举）：

- **儿童与家庭管理局（ACF）**：负责管理 60 多个项目，为家庭和儿童提供福利与服务，包括促进经济和社会福祉。ACF 在 HHS 人工智能战略计划中的职责将聚焦于确保向儿童和家庭有效且公平地提供服务，以促进最佳健康状况。
- **医疗保健研究与质量局（AHRQ）**：致力于通过研究、技术评估以及传播与实施工作，提高全体美国人医疗保健的质量、安全性、效率和有效性。AHRQ 在 HHS 人工智能战略计划中的职责将集中在推动和开展关于安全使用人工智能以实现高质量医疗服务的研究，传播可行的、基于证据的人工智能知识，以及提供覆盖决策所需的证据。
- **高级卫生研究计划局（ARPA - H）**：在多个重点领域开展变革性、高影响力的医疗保健研究，包括推进技术解决方案、打造有韧性的健康生态系统以及推动可扩展的解决方案。ARPA - H 在 HHS 人工智能战略计划中的职责将着重于发放奖励，以催化能够改善医疗服务提供的前沿研究。
- **疾病控制与预防中心（CDC）**：提供关于主要疾病医疗服务提供的指南和研究，支持公共卫生项目资金投入，并可能利用人工智能为医疗服务提供提供信息和支持。CDC 在 HHS 人工智能战略计划中的职责将聚焦于研究人工智能在疾病预防方面的功效，并在公共卫生工作中实施人工智能。
- **医疗保险和医疗补助服务中心（CMS）**：管理主要的公共医疗保健支付方项目（如医疗保险和医疗补助），并可能参与制定特定项目或服务的支付和覆盖政策。CMS 在 HHS 人工智能战略计划中的职责将集中在确定人工智能赋能的医疗保健服务的覆盖范围和支付方式，监督和认证州信息技术系统及数据收集标准，并向医疗服务提供者、各州及其他利益相关者提供技术援助。在适当情况下，CMS 将寻求利用支付和监管政策，确保支付方和医疗服务提供者以可信赖、负责任的方式使用人工智能。

- **食品药品监督管理局 (FDA)**：协助确保人用和兽用药品、生物制品以及医疗器械对于其预期用途而言是安全有效的，并且确保发射辐射的电子产品是安全的。随着人工智能在医疗产品及其研发、制造运营和使用中变得愈发重要，FDA 将在监管和支持利益相关者方面继续发挥作用。
- **卫生资源与服务管理局 (HRSA)**：为全国需求最迫切的社区提供公平的医疗保健服务，包括通过支持低收入人群、艾滋病毒感染者、孕妇、儿童、父母、农村社区、移植患者以及卫生人力的项目来实现。HRSA 在 HHS 人工智能战略计划中的职责将聚焦于确保公平使用人工智能，使服务不足的社区受益，并对未来几代医疗保健专业人员进行教育和培训。
- **印第安健康服务局 (IHS)**：为美国印第安人和阿拉斯加原住民社区提供医疗保健服务。IHS 在 HHS 人工智能战略计划中的职责将聚焦于在这些人群中实施医疗服务提供，并确保人工智能指南适用于相关的州、部落、地方和领地 (STLT)。
- **国立卫生研究院 (NIH)**：支持并在美国及国外开展生物医学和行为研究，并可通过其倡议帮助对卫生人力进行人工智能教育并促进创新。NIH 在 HHS 人工智能战略计划中的职责将聚焦于支持关于人工智能对生物医学和行为健康影响的研究，基于研究在这些领域制定标准，并释放资金以促进 HHS 各服务领域负责任地使用人工智能。
- **药物滥用和精神健康服务管理局 (SAMHSA)**：引领公共卫生工作，以推进国家的行为健康，改善患有精神和物质使用障碍的个人及其家庭的生活。SAMHSA 在 HHS 人工智能战略计划中的职责将聚焦于向 STLT 社区提供赠款资金和指南，并收集、分析和分发行为健康数据，以评估项目、改进政策，并提高对预防、减少伤害、治疗和康复资源的认识。

其他联邦机构：

HHS 还与许多其他联邦部门密切合作，如退伍军人事务部和住房与城市发展部。

患者、受益人和他们的护理人员：

初级医疗服务接受者将以某种患者身份与医疗保健系统互动；2020 年，83.4% 的成年人和 94.0% 的儿童报告称他们在前一年曾就医或看其他医疗保健提供者。护理人员，有时担任监护人，在为婴儿、儿童、青少年和老年家庭成员提供护理方面也发挥着关键作用。

医疗服务提供者：

这些是美国提供医疗服务的主要载体，包括：

- **医疗保健机构和系统**：美国医疗系统除了一系列急性后期护理机构、门诊诊所和长期护理机构外，还包括约 6100 家医院（从小型社区组织到全国性系统）。

- **临床医生和支持人员**：2022 年，美国约有 1500 万临床工作人员，包括 93.3 万名在职医生、340 万名注册护士和 140 万名个人护理助理，以及其他临床工作人员（如专科医生、助理、治疗师和技术人员）。
- **非临床工作人员**：非临床工作人员在组织和提供医疗保健方面发挥着关键作用（如供应链、维护、接待、人力资源和财务、通信和信息技术），并且在行政环境中也可能接触到人工智能赋能的工具。
- **医疗保健行政管理人员**：医疗和健康服务管理人员协助协调和监督医疗服务提供组织的复杂运营；2023 年，美国有 567,200 名医疗保健行政管理人员。此外，高级管理人员、受托人和董事会推动医疗服务提供组织的总体战略，并就人工智能投资做出决策。

支付方：

这些是为患者护理提供资金，并根据患者需求帮助将患者与合适的医疗服务提供者和服务相连接的公共和私人组织，包括：

- **公共支付方（如州医疗补助和其他政府机构）**：支持实施监管、融资和服务提供的机构。
- **私人支付方**：支持融资和护理的全国性、地区性和地方性支付方。
- **雇主**：雇主赞助的医疗保险在美国管理式医疗覆盖人口中占 54%（通常由私人支付方管理）。雇主积极关注确保向其员工提供的医疗服务的质量和安全性。

州、部落、地方和领地（STLT）政府：

这些实体直接开展各种医疗服务提供活动，包括提供护理、融资以及对私营和公共部门活动进行监管监督。

其他支持医疗服务提供的实体：

- **科技公司**：各种技术供应商积极为医疗保健机构开发技术，从多元化的大型科技公司到专注于医疗保健服务和技术的供应商，如电子健康记录（EHR）、收入周期管理（RCM）及其他辅助服务供应商。
- **研究机构**：学术研究机构与医疗保健机构合作，推动新发现，解锁有可能改变医疗标准的新治疗方式（如增强患者服务、新的临床创新、缓解质量和安全问题、新设计的组织工作流程）。
- **生物制药和医疗器械公司**：虽然本计划其他章节涵盖了包括药品、生物制品和器械在内的医疗产品研发和发现的具体内容，但这些组织也通过对临床环境中部署的人工智能进行上市后监测、维护和监督，参与医疗服务提供。
- **非营利组织和社区组织（CBOs）**：这些实体中的许多都支持直接提供转诊和护理协调服务。

3.3 人工智能在医疗保健交付中的应用机会

人工智能有潜力改变护理交付过程，但它也带来了必须加以监测的固有风险，以确保对患者产生积极影响和安全。人工智能可以支持医疗保健系统的五种方式包括：

- 1. 提高患者护理的质量和安全性：**医疗错误，包括错误和 / 或延迟诊断，可能导致不良患者结果。人工智能有潜力通过快速处理大量不同信息、检测人类观察不易察觉的模式并引导临床医生做出更有可能的诊断来加速诊断并预防不良事件。人工智能还可以增强护理模式和卫生服务研究，以开发更好地使临床医生、支付者和患者受益的创新。
- 2. 改善患者体验：**人工智能有潜力通过更高效和量身定制的服务来提高患者满意度，这些服务更好地满足他们的需求。人工智能还可以为患者提供工具，以更好地理解医疗信息，包括他们自己的病历和健康状况，并促进与提供者和支付者更积极的沟通（例如通过共享可解释和相关的面向患者的信息）。
- 3. 自动化行政流程并减轻劳动力负担和倦怠：**医疗保健交付行政复杂性的增长，加上医疗保健劳动力短缺，特别是在初级保健方面，加剧了这些已经要求很高的工作环境中的倦怠。人工智能在行政环境中的应用 —— 包括文档记录、成员 / 患者通信和索赔处理 —— 可以节省资源，并为组织提供更多带宽来加强护理交付。
- 4. 增强公平性和可及性：**医疗保健差异在医疗保健中持续存在，结果可能因社会经济地位、地理位置、人口因素等而异。人们越来越意识到健康的社会决定因素和与健康相关的社会需求对健康结果的重要性。人工智能系统有能力纳入社会决定因素和其他信息，以识别高危患者、以患者首选的语言和识字水平进行沟通、克服残疾人的获取障碍，并推荐更适合个人情况的服务和资源。
- 5. 降低成本曲线：**美国仍然是全球成本最高的医疗保健系统，这限制了美国人获得护理的机会，并阻碍了美国的经济生产力。总体而言，在医疗保健交付价值链中使用人工智能可以减少行政开销、提高资产和资源利用率并减少不良事件，一些报告估计这可能使年度国家医疗保健支出减少多达 10%。

3.4 人工智能在医疗保健交付中的趋势

当前趋势表明，人工智能在医疗保健交付中的创新应用正在迅速发展。然而，其使用仍然存在障碍。关键趋势包括：

- 1. 对健康人工智能的投资规模大且不断增长：**自 2021 年以来，人工智能占有所有医疗保健风险投资资金的 25%，总计超过 190 亿美元。根据初步报告，大约三分之二的投资进入了人工智能的临床应用领域，另外三分之一用于行政用途。
- 2. 对医疗保健交付中使用人工智能的热情和担忧并存：**最近一项对 100 名医疗保健高管的调查显示，超过 70% 的人已经在追求或实施该技术。然而，在另一项调查中，约 40% 的医生表示他们对使用人工智能既热情又担忧。有人担心人工智能的使用可能会导

致医疗工作岗位格局的变化，并影响医患关系。患者对人工智能也有类似的担忧，一项调查结果显示，约 60% 的受访者对医疗服务提供者依赖人工智能感到不安。本章“行动计划”部分对这些风险及相关缓解措施有进一步讨论。

3. 不同医疗学科对人工智能的使用存在差异：在 20 世纪 90 年代，早期的机器学习应用于医疗数据，开发出了第一批基于机器学习的诊断系统。如今，人工智能创新持续进行，临床决策支持已成为现代临床工作流程的关键工具。目前，某些人工智能和机器学习应用——特别是在放射学领域（如审查心电图、磁共振成像扫描和皮肤图像等医学图像类型）——已被广泛接受。虽然放射学领域的人工智能应用已经成熟，但在其他学科，如病理学、心脏病学和初级保健领域，人工智能的使用仍在增长。本章“用例和风险”部分对这些用例有进一步分析。

4. 行政人工智能应用的创新和使用增加：鉴于与临床用例相比开发成本较低，以及生成式人工智能和大型语言模型技术的出现，人工智能在行政任务中的应用在过去几年有所进展。最近的应用包括“从医生笔记中提取药物名称、回复患者行政问题、总结医疗对话以及撰写病史和身体评估”。根据美国医学协会的一项调查，54% 的医生对在其实践中使用人工智能（特别是在行政任务如文档记录和图表绘制方面）充满热情。

5. 各组织的数据和技术系统存在异质性：医疗保健组织在使用人工智能所需的技术和资源（包括数据管理、临床和行政应用以及核心基础设施如云计算）方面的获取差异，影响了当前的使用情况。医疗保健系统和电子健康记录中数据模式（如数字、文本、图像、视频和音频）和标准的异质性，为人工智能在整个医疗保健领域的应用带来了额外障碍。这种异质性也影响了组织在决定构建、合作（如与人工智能供应商）或从其他方采购何种解决方案以及采购程度（如人工智能、生成式人工智能或非人工智能干预措施）方面的决策。

3.5 医疗保健交付中人工智能的潜在用例和风险

医疗保健交付和融资包括广泛的活动，所有这些活动都可能受到现有和新兴人工智能的影响，尽管有些可能受到的影响更大。本章所考虑的医疗保健交付和融资中人工智能的使用，可以从医疗保健交付（如诊断服务、患者护理交付）、融资（如索赔处理、提供者网络管理）和研究的价值链活动角度来看待（见展览 8）。

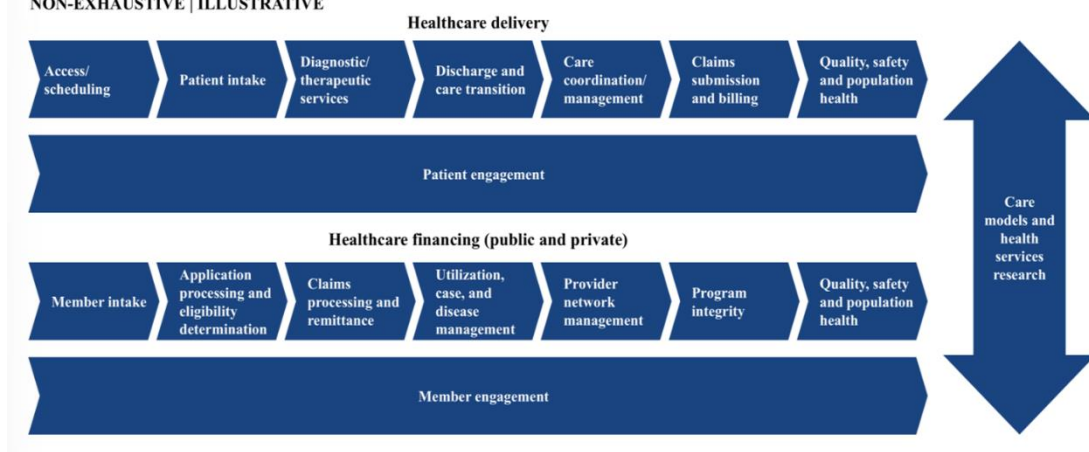
人工智能用例在技术类型和复杂性方面存在差异（如简单的基于规则的自动化与复杂的大型语言模型），因此，一些用例在整个医疗保健系统中的使用率高于其他处于较早期测试阶段的用例。

医疗保健交付中人工智能也带来了广泛的风险，包括对患者安全的影响、医患关系的恶化以及由算法偏差导致的护理障碍或不当管理。如前文所述，HHS 及其部门（如 CMS）提供了考虑和缓解医疗保健人工智能风险的框架，如 FAVES。

图 8：医疗保健交付和融资价值链

Exhibit 8: Healthcare Delivery and Financing Value Chains

NON-EXHAUSTIVE | ILLUSTRATIVE



3.5.1 交付中的人工智能

虽然提供者组织提供的个别活动在规模和重点上会有很大差异（如初级保健诊所、大型多专科集团、学术医疗中心、州机构和联邦合格健康中心），但价值链旨在描述医疗保健交付的核心功能集以及人工智能的潜在益处或应用。

人工智能在价值链中的创新、开发和使用情况并不一致——在行政职能方面（包括具有临床和非临床影响的职能，如手术室优化、呼叫中心启用、人才管理和后台管理）相对更先进，而在诊断和治疗服务方面的人工智能应用在放射学之外仍不常见。

总体而言，在数据容易获取的用例中（如通过电子健康记录或可穿戴设备），人工智能的使用率相对较高，而在数据有限的复杂病例应用中（即由于模型不准确或对特定人群存在偏差的风险）仍处于起步阶段。需要连接不同数据源（如远程监测以及医院和家庭护理记录）的领域，如护理协调和过渡，可能存在机会，但由于连接基础数据的挑战，其使用仍然有限。

较大的医院在人工智能使用方面进展较快，而较小的医院和医生团体则处于人工智能应用的早期阶段，正在试点一些用例。然而，如前文所述，某些人工智能用例的相对价值将根据个别组织的特征（如提供者规模、需求、资源、现有能力和服务区域）而有所不同。

在下表中，HHS 重点列出了医疗保健交付价值链中人工智能的潜在益处和风险的详尽清单。请注意，以下用例突出了该领域各种利益相关者使用人工智能的现有或潜在方式。有关 HHS 及其部门如何使用人工智能的详细信息，请参考 HHS 人工智能用例清单 2024。此外，融资和研究中的用例和风险分别在 3.5.2 融资中的人工智能和 3.5.3 护理模式和卫生服务研究中的人工智能中进行讨论。

- **功能组件 1：预约和 / 或调度**
 - 安排患者预约和服务的过程

潜在用例（非详尽）	潜在风险（非详尽）
简化和自动化的调度工具以优化效率 例如，预测分析以减少爽约情况 有针对性的干预措施（如外联）可以大幅提高最有可能错过预约的患者的就诊率 例如，预约安排优化 人工智能可以通过预测患者的预约偏好和可预约时间来优化安排，减少等待时间，提高诊所效率。 例如，手术室调度优化 人工智能可以分析手术安排、患者数据和资源可用性，以优化手术室使用，减少闲置时间并提高手术通量。	可能引入偏差 例如，不匹配的超额预订 对所有预约应用一刀切的超额预订方法，根据爽约率可能会不成比例地影响具有某些特征（如社会经济地位、交通不便和对医生或医院的恐惧）的患者 例如，预约调度优化 人工智能可以通过预测患者预约偏好和可用性、减少等待时间和提高诊所效率来优化调度 例如，手术室调度优化 人工智能可以分析手术时间表、患者数据和资源可用性，以优化手术室使用，减少停机时间并提高手术吞吐量 例如，过度强调加剧调度差异的变量 如果基于当前资源分配数据训练的算法过度重视某些变量（如手术盈利能力、保险类型），用于手术调度（如手术室调度）的人工智能使用可能会加剧护理获取的差异

- **功能组件 2：患者接待和支持**

- 收集和验证患者信息（包括病史和保险详细信息）以准备治疗并确保顺利行政流程的初始阶段

潜在用例（非详尽）	潜在风险（非详尽）
个性化的人工智能辅助患者接待流程以提高效率和患者满意度 例如，简化的患者数据收集 自动生成和跟踪发送给患者的通信，以最大限度地减少重复数据收集和患者负担 例如，自动化人工智能语音技术 人工智能驱动的对话式语音技术可自动执行患者接待流程（如通过录音和转	可能加剧患者信任担忧 例如，过度收集患者数据 为人工智能模型过度收集数据（或对数据滥用的感知，即使不准确）可能会导致患者在护理交付过程中感到不适，并产生或加剧不信任，特别是对于那些可能已经对医疗保健系统有负面看法的人群

潜在用例（非详尽）	潜在风险（非详尽）
录)	
自动化工具可减少行政任务, 让工作人员腾出精力专注于患者护理及更复杂的问题。 例如, 优化患者请求处理 人工智能虚拟助手能够快速回应患者的简单请求 (例如, 在新冠疫情期间, 某医疗系统利用自然语言处理驱动的聊天机器人, 将大量涌入的患者来电引导至合适的系统, 以协助处理他们的请求)。	可能阻碍患者获得医疗服务 例如, 基于患者数据的人工智能做出错误决策。人工智能收集到的错误数据可能导致不恰当的决策以及拒绝提供服务。

- **功能组件 3：诊断 / 治疗服务**

- 由先进系统（如电子健康记录、临床决策支持、可穿戴设备和远程医疗工具）支持的医疗护理（包括诊断和治疗）的提供，以提高护理的质量和效率

潜在用例（非详尽）	潜在风险（非详尽）
患者信息的自动化文档记录和总结以提高医护人员效率 例如, 环境监听 人工智能驱动的环境监听系统可以实时捕获和转录患者 - 提供者互动, 促进更准确的文档记录和诊断, 并使提供者能够更多地关注患者护理并改善患者体验	不适当应用的可能性 例如, 虚构信息 自动化文档系统可能会生成关于患者病史的虚假信息, 并导致不适当的护理建议, 凸显了人工监督和强大的虚构信息检测方法的重要性 例如, 人工智能影响医患关系和信任 患者表示担心使用人工智能进行临床决策可能会恶化医患关系, 因为人工智能不断自动化通常由

潜在用例（非详尽）	潜在风险（非详尽）
	人类完成的任务，特别是考虑到经历医疗状况的情感和个人性质，凸显了在医疗保健交付中富有同情心和同理心互动的重要性
自动化智能工具支持诊断和治疗方案评估并揭示关于患者病情的关键见解 示例：预测和风险识别：人工智能算法可以分析患者健康指标来预测疾病结果（例如，某医疗系统使用人工智能算法，通过结合电子健康记录（EHR）数据与血压和心率测量值来预测患者的脓毒症）。 示例：精准医疗：人工智能赋能的临床决策支持（CDS）工具可帮助医生考虑最佳干预措施，并帮助揭示关于患者病情的关键（且可能难以追溯）见解	潜在的对健康数据的误用或误解 示例：人工智能提供的无效治疗计划：人工智能生成的行为健康治疗决策支持可能导致误诊和治疗，从而可能恶化行为健康结果和信任。
分析患者数据以制定针对性干预措施或教育材料 示例：通过多种数据格式进行情感分析：人工智能可以处理非结构化数据（例如，社交媒体上发布的文本、用户输入），生成关于心理健康的观点摘要。这些可用于开发和传播个性化的教育材料、指南、策略和转诊。 示例：人工智能分析以制定综合干预措施：人工智能可以帮助创建整体治疗计划，将多种类型的干预措施（例如，行为和临床干预）相结合，如引导饮食监测和与 CDS 相结合的人工智能定制教育（例如，健康指标测量和个性化用药计划），用于糖尿病患者。 示例：提供提醒和测量药物摄入的人工智能技术：人工智能工具，如智能手机应用程序，可以通过日常监测和提醒来评估和鼓励依	示例：健康技术可能未考虑个体的细微差别：人工智能工具可能未考虑人口统计学和社会决定因素（SDOH）等因素，例如沟通障碍，这可能会增加患者的技术担忧，导致患者满意度、信任度和护理有效性降低。

潜在用例（非详尽）	潜在风险（非详尽）
从性（例如，使用智能手机摄像头确认药物摄入）。	

- **功能组件 4：出院和护理过渡**

- 通过集成系统和患者参与平台管理患者从一种护理环境转移到另一种护理环境的过程，确保护理的连续性和适当的随访

潜在用例（非详尽）	潜在风险（非详尽）
分析患者情况并实现更个性化和高效护理过渡的人工智能算法 例如，面向患者的虚拟护理助手 人工智能可以通过解释诊断和护理管理计划来增加教育和透明度，为患者提供 24/7 的资源，通过虚拟护理助手或聊天机器人教育他们并提供及时信息 例如，尽量减少与临床医生潜在接触的聊天机器人 聊天机器人可以帮助鼓励和为那些可能因尴尬或污名化而不愿进行面对面咨询的患者提供护理	不适当应用的可能性 例如，不适当建议的虚构信息 人工智能模型可能会在数据分析中出错，错误转录记录或向临床医生传达虚假信息 例如，关键技能退化 额外引入人工智能工具可能会导致临床医生过度依赖这些技术，可能导致在健康的细微领域技能退化，特别是在人类同理心和参与发挥重要作用的地方

- **功能组件 5：护理协调和管理**

- 利用数字工具和分析对不同服务和提供者的患者护理进行持续管理，以增强护理协调、患者参与和整体健康结果

潜在用例（非详尽）	潜在风险（非详尽）
远程监测患者状况以提高患者护理的有效性和及时性	可能引入偏差 例如，按人口统计学错误分层风险

潜在用例（非详尽）	潜在风险（非详尽）
例如，慢性病管理 人工智能决策辅助工具可以通过为患者提供支持提醒、预测问题并向提供者和患者标记护理需求的工具，支持持续疾病管理。此外，它们可以在医院环境中用于跨疾病领域监测护理（如糖尿病患者因其他急性需求住院时的血糖变化）	用于护理协调决策的人工智能算法可能容易受到偏差影响，因为它们可能会将相同的风险级别分配给具有不同特征的患者，而这些特征在确定风险时应予以考虑（如一个健康系统使用的人工智能算法减少了被确定需要护理的少数族裔患者数量，尽管该群体患者病情更严重且需要更多护理）

- **功能组件 6：索赔提交和计费**

- 提交报销索赔和管理计费通常是自动化的，以确保及时准确的支付、减少拒付并优化收入周期管理

潜在益处和示例用例（非详尽）	潜在风险（非详尽）
帮助测量和协助医生选择和记录最佳干预措施的工具 例如，计费代码自动化和分析 根据非结构化笔记和数据自动生成计费代码并检查计费准确性	患者护理障碍增加的可能性 例如，不准确的索赔提交 由于模型故障（如不良 / 暴露的数据、分析方法、解释）导致的人工智能不准确索赔提交可能会增加医疗专业人员责任和罚款 例如，由于支付完整性 / 收入周期管理竞争导致成本增加 随着提供者投资人工智能以优化收入，支付者投资人工智能以增加支付完整性，系统的潜在成本增加 —— 增加了影响患者的额外风险（如在支付者拒付的情况下医院向患者计费）

- **功能组件 7：质量、安全和人群健康**

- 确保医疗服务符合既定的质量和安全标准，使用人工智能支持工具、决策支持系统和持续监测来改善临床护理和组织绩效

潜在用例（非详尽）	潜在风险（非详尽）
增强患者护理和医院质量措施的人工智能工具 例如，不良事件和再入院预防 人工智能可以远程监测患者状况，通过识别潜在恶化风险并优先安排干预措施来预防再入院，确保及时有效的护理 例如，质量测量 用于更准确和高效医院质量测量的抽象和分析工具	可能存在偏差 例如，某些人群在训练数据中代表性不足 由于历史上护理交付的获取和质量差异，基础训练数据可能存在偏差

3.5.2 融资中的人工智能

融资领域有各种各样的支付者（如医疗保险、州医疗补助机构、大型全国性保险公司、区域专科支付者和管理式医疗组织）。这些组织在服务人群和支付结构（如基于价值的护理、按服务收费）方面存在关键差异。在这些方面，对于这些支付者存在广泛的用例和风险，包括下表中列出的示例。

在融资中，人工智能和大型语言模型越来越多地用于一系列功能和任务，包括预先授权、临床审查评估、利用管理和索赔裁决。鉴于该行业广泛的数据分析和文档处理，预计未来几年将出现大量新的用例。这一领域的扩展并非没有挑战：国会一直在就使用人工智能和算法拒绝预先授权请求进行诉讼和表示担忧，特别是在人工智能如何遵守影响支付者决策的州和联邦法规方面。

功能组件 1：成员登记

将个人纳入医疗保险计划的流程，确保准确获取并维护他们的信息，以便用于未来的交互。

潜在用例（非详尽）	潜在风险（非详尽）
人工智能可大幅减少以手动申请为中心的工作量 示例：申请审核：人工智能可支持快速审核申请，以识别缺失信息，检查其他资格（例如，二次保险覆盖），并支持其他功能。 示例：自适应面向客户的聊天机器	可能引入偏见 示例：错误地拒绝资格：如果没有适当的校准或人工干预，模型可能会拒绝资格，特别是对于那些在历史上保险覆盖情况

潜在用例（非详尽）	潜在风险（非详尽）
人：由人工智能驱动的聊天机器人可经过训练，处理各种各样的查询 — — 从资格问题到申请状态更新 — — 为会员提供即时且准确的回复，减轻呼叫中心的负担。	更为复杂的人群，这会给患者接受必要治疗程序造成重大障碍。 示例：加剧服务不足人群对医疗服务的不信任：对那些本就不太可能寻求医疗服务和支持的人群给出不准确的回复，可能会进一步抑制他们寻求医疗服务的行为。

- **功能组件 2：申请处理和资格确定**

- 审查和验证申请以确定申请人是否符合覆盖标准，确保只有符合条件的个人获得福利

潜在用例（非详尽）	潜在风险（非详尽）
人工智能显著减少手动申请处理工作量 例如，申请审查 人工智能可以支持快速审查申请，识别缺失信息，检查其他资格（如二次保险）并支持其他功能 例如，自适应客户聊天机器人 人工智能驱动的聊天机器人可以被训练来处理各种查询 —— 从资格问题到申请状态更新 —— 为成员提供即时准确的回复并减少呼叫中心负担	可能引入偏差 例如，错误拒绝资格 如果没有适当校准或人工监督，模型可能会错误地拒绝资格 —— 特别是对于历史上保险覆盖情况更复杂的人群 —— 给患者获得必要程序带来重大障碍 例如，加剧服务不足人群对护理的不信任 对已经不太可能寻求护理和支持的人群的不准确回复可能会进一步阻碍他们寻求护理的行为

- **功能组件 3：索赔处理和支付**

- 处理和裁决医疗服务提供者提交的索赔，根据政策条款和提供的服务确保及时支付或拒付

潜在用例（非详尽）	潜在风险（非详尽）
自动人工智能处理复杂索赔数据以简化决策 例如，快速跟踪索赔批准 预测分析可以生成摘要并快速评估索赔的有效性以快速批准 例如，自动化索赔审查 人工智能算法可以自动审查索赔中的错误、不一致和政策条款合规性，加快流程并减少手动工作量	可能加剧系统成本 例如，由于支付完整性 / 收入周期管理竞争导致成本增加 随着提供者投资人工智能以优化收入，支付者投资人工智能工具以增加支付完整性，通过行政浪费可能会产生增量成本——这可能会影响患者（如在支付者拒付的情况下医院向患者计费）

• **功能组件 4：利用、案例和疾病管理**

- 监测和管理医疗服务的使用，确保其必要且具有成本效益，从而优化资源使用并控制成本

潜在用例（非详尽）	潜在风险（非详尽）
预测分析工具优化向患者提供的医疗服务 例如，患者再入院分析和预防 人工智能干预措施可以支持病例管理计划，通过预测哪些患者处于更高风险并支持有针对性的干预措施来防止未来再入院（在一个示例中，人工智能干预措施将再入院率降低了 55%） 自动人工智能处理利用管理和预先授权 例如，预先授权裁决 人工智能可以简化预先授权流程，快速验证必要的医疗信息并自动化批准工作流程，减少患者护理延迟	可能产生不适当结果 例如，用于覆盖范围确定的人工智能决策支持工具 用于支持服务覆盖范围确定的人工智能决策支持工具可能与覆盖条款、特定患者情况不一致，或不遵守适用的联邦或州法律

功能组件 5：提供商网络管理

管理与医疗保健提供商的关系和合同，以确保为成员建立一个强大且有效的网络，促进获得必要服务

潜在用例（非详尽）	潜在风险（非详尽）
将定量网络指标（例如，费率、资质合规性）进行比较的算法，以确定可变性领域和标准化潜力 例如，提供商费率比较 人工智能可以比较不同提供商和服务的费率，帮助支付方和患者就具有成本效益的护理选项做出明智决策，并协商更好的费率。<sup>453</sup>	可能加剧偏见 例如，由人工智能算法导致的偏见增加安全网医院，通常利润率较低且为未得到充分代表的人群提供护理，在支付方使用复杂的人工智能算法管理其网络策略时，可能在与支付方的谈判中处于更不利的地位。<sup>455, 456</sup>
简化文档流程以支持扩展提供商网络的算法 例如，自动化提供商资质认证 人工智能可以自动化提供商资质验证，确保所有必要的资格和认证都是最新的，并减轻医疗保健组织的行政负担。	

- 功能组件 6：计划完整性

- 实施措施，包括高级分析，以防止医疗保健保险系统内的欺诈、浪费和滥用，确保计划的完整性和可持续性

潜在用例（非详尽）	潜在风险（非详尽）
用于检测和减轻欺诈以保护患者的算法 示例：欺诈检测：医疗服务提供者排名算法可以利用公开和私有可用数据的语料库来识别欺诈行为。医疗保险和医疗补助服务中心（CMS）已推出一个欺诈预防系统，该系统使用预测分析在付款前筛查理赔申请，利用标记欺诈的指标并实施保护性干预措施。	可能产生意外后果或不当结果 示例：加重经济负担：支付方对人工智能工具进行投资，若这些工具导致更多不利的承保决定，可能会对患者和医疗机构造成经济影响。因为医院为弥补因更多拒赔而损失的收入，会提高收费。

- 功能组件 7：质量、安全和人群健康

- 确保向成员提供的医疗服务符合既定的质量和安全标准，包括持续监测和改进这些标准

潜在用例（非详尽）	潜在风险（非详尽）
通过人工智能模型监测患者和医疗服务提供者的各项指标(如情感分析)，收集反馈以提升医疗服务质量。 例如，患者体验分析：人工智能能够分析来自调查问卷及其他渠道的患者反馈，从而明确患者满意度、发掘趋势，并找出医疗服务中有待改进的方面，为支付方就其医疗服务提供网络内的服务质量提供切实可行的见解。	存在引入偏差的可能性 示例：无法识别特定患者群体的疾病：基于特定患者群体训练的人工智能算法可能存在偏差，从而导致关于患者安全的结论不准确。例如，基于某医院电子健康记录（EHR）构建的脓毒症预测算法，仅能识别 7% 患者群体中的该病症，这就延误了其他有需要患者的治疗，同时也未能准确反映患者护理质量。

3.5.3 护理模式和卫生服务研究中的人工智能

在应用研究环境中，人工智能在护理模式和卫生服务开发中的应用正在增长。人工智能也为非设备行为干预措施（如认知行为疗法、营养咨询）的开发提供信息，这可能导致现有干预措施的生成、修改、适应或完善。HHS 部门支持这些领域的研究和创新，包括 AHRQ（如人工智能与安全 NOFO、减轻算法偏差指南）、CMS（如 CMMI 成果挑战）、NIH（如行为和社会科学研究办公室）和 SAMHSA（如行为健康统计和质量中心）。

护理模式和卫生服务研究

通过创新的、数据驱动的和基于价值的方法分析和优化医疗保健交付、劳动力模型、财务绩效和患者结果，以提高卫生系统绩效和公平性。

请注意：这是指基于人工智能的护理模式研究，可能涉及医疗产品，但主要涉及使用人工智能改善医疗保健交付。有关医疗产品开发的讨论见其他章节，特别是“医疗产品开发、安全性和有效性”。

潜在用例（非详尽列举）	潜在风险（非详尽列举）
临床路径与护理模式优化 / 生成 例如：人工智能生成的护理路径 人工智能可以推荐和优化临床护理路径，确保患者护理符合循证指南，并减少从业者之间临床护理的差异。 例如：基于人群数据强化的护理模式 人工智能可以综合大量数据并生成定制化护理模式。通过围绕患者治疗结果调整激励措施，人工智能可以帮助支付方开发基于价值的护理模式。预测分析可以识别趋势并帮助为患者开发护理模式。 例如：用于测量患者状况的数字孪生 人工智能可以分析来自各种来源（包括电子健康记录、可穿戴设备、医疗设备等）的患者数据，生成数字孪生，有助于早期发现健康风险并创建主动干预措施。	
为患者治疗结果和护理模式研究提供信息的预测工具 示例：支持服药依从性的人工智能智能手机应用程序：一款具备人工智能功能的智能手机应用程序可以提供服药提醒和剂量说明，然后确认药物摄入情况，以检测不依从行为并预测未来的不依从情况。这些数据以及由此产生的任何研究结果可用于为直接护理和护理模式设计提供信息。 示例：用于医院内心脏骤停的预测性快速响应系统：基于人工智能的算法可用于预测病情恶化事件（如心脏骤停和意外入住重症监护病房），这可用于改进决策制定和护理模式设计。	可能引入偏差 例如：针对患者群体研究得出不准确结论 基于特定患者群体训练的人工智能算法可能存在偏差，导致研究结果具有误导性，无法平等适用于各个群体，或者延续现有偏差。

3.6 行动计划

鉴于医疗保健交付中不断变化的人工智能格局，HHS 已经采取了多项措施，包括发布新的指南和规则以及启动与健康人工智能相关的计划，以促进负责任的人工智能发展。以下行动计划遵循支持 HHS 人工智能战略的四个目标：1. 促进健康人工智能创新和使用；2. 促进可信人工智能开发和道德及负责任使用；3. 使人工智能技术和资源民主化；4. 培养人工智能赋能的劳动力和组织文化。对于每个目标，行动计划提供了背景、HHS 和其他相关联邦行动迄今为止的概述，以及 HHS 将采取的近期和长期优先事项。HHS 认识

到，随着技术的发展，该行动计划需要不断修订，并致力于提供结构和灵活性，以确保产生长期影响。

3.6.1 促进健康人工智能创新和使用

HHS 有机会通过以下行动安全地增加人工智能创新和使用：

1. 支持收集人工智能干预措施的有效性、安全性和风险缓解证据以及在医疗保健交付环境中实施的最佳实践
2. 提供关于监督、医疗责任以及隐私和安全保护的指南和资源，以增强组织开发人工智能的信心
3. 确保人工智能开发者和潜在部署者清楚覆盖和支付确定流程，以鼓励人工智能开发

以下，HHS 讨论了背景、HHS 和其他联邦行动迄今为止的情况，以及促进医疗保健交付中健康人工智能创新和使用的计划。

1. 支持收集人工智能干预措施的有效性、安全性和风险缓解证据以及在医疗保健交付环境中实施的最佳实践

背景：

在医疗保健交付环境中，临床医生和其他领导者对人工智能的信心、理解 and 对其潜在影响的担忧存在差异。一些学科，如放射学，在临床环境中使用人工智能有更成熟的记录。相比之下，在目前的情况下，其他学科在行政环境之外不太可能看到人工智能应用。根据美国医学协会的一项调查，56% 的医生认为最有前途的人工智能用例是支持行政任务。对人工智能在复杂临床环境中的应用进行进一步研究可以通过为人工智能的适当和安全使用提供基于证据的基础来解锁创新并激励使用。这些努力还可以旨在建立证据来解决临床医生和其他利益相关者的担忧，以确保人工智能以对患者和参与其护理的人最有帮助的方式被使用。通过基于对其益处的充分了解建立对这些技术的信心，这种方法还将有助于维持人工智能的有效和负责任使用。

鉴于美国医疗保健组织在人工智能准备程度和基础设施方面高度多样化，额外的资源、指南和教育也将帮助组织评估对人工智能的投资决策。

HHS 和其他联邦行动迄今为止（非详尽）：

- ASTP LEAP 在健康信息技术合作协议奖项为患者护理的人工智能解决方案的高级开发提供了资金机会。
- CMS 人工智能健康成果挑战为创新者提供了机会，展示他们可用于预测医疗保险受益人的患者健康结果的人工智能工具，这些工具可能用于 CMS 的创新支付和服务交付模型。
- 在 COVID - 19 大流行期间，NIH 的 COVID - 19 医学成像参与了一项多机构努力，利用医学成像技术筛查感染的心脏和肺特征，以评估疾病严重程度并提出治疗建议。

- 国家心理健康研究所（NIMH）的数字全球心理健康计划资助关于适合低收入和中等收入国家的数字心理健康技术的开发、测试、实施和成本效益的研究。它强调利用人工智能和 / 或其他新颖的计算和统计方法来改善心理健康的预防、诊断和治疗，贯穿治疗轨迹和连续护理。
- 总务管理局（GSA）的技术转型服务（TTS）和其他计划支持医疗保健交付研究，包括通过技术提升和创新支持，并可扩展到包括人工智能。
- SAMHSA 的《利用技术增强危机连续体中的服务获取》出版物强调了帮助有需要的人获得关键服务的技术创新用途，包括人工智能如何帮助疾病筛查和交付（如个性化自助心理健康应用程序）。
- AHRQ 人工智能与医疗保健安全 NOFO 邀请资助申请，通过确定（1）某些突破性的人工智能系统使用是否以及如何影响患者安全，以及（2）如何安全地实施和使用人工智能系统来支持医疗保健安全。
- SAMHSA 神经网络分析利用人工智能神经网络分析物质使用问题、焦虑障碍和抑郁障碍的共现情况。研究结果表明，心理健康诊所应提供综合治疗计划并筛查各种情况和因素。

HHS 近期优先事项：

- 支持关于在医疗保健交付环境中采购、部署和监测人工智能工具的最佳实践的卫生服务研究（如 AHRQ 医疗保健安全和人工智能 NOFO）。
- 在现有推动医疗保健交付中人工智能创新的“挑战”倡议基础上再接再厉，如 CMS 人工智能健康成果挑战和 NIH CRDC 人工智能数据准备（AIDR）挑战。
- 探索机会扩展促进医疗保健交付背景下人工智能创新的倡议，如 GSA 的 TTS。
- 提供关于如何在医疗机构内测试和试点人工智能应用，然后在护理交付中全面实施指南。

2. 提供关于监督、医疗责任以及隐私和安全保护的指南和资源，以增强组织开发和部署人工智能的信心

背景：

在不知道新的人工智能干预措施是否已被适当实体审查，或者这些实体是否考虑了患者结果、安全、隐私和其他因素的情况下，提供者不愿意部署这些措施。在不清楚使用这些工具可能带来的潜在责任的情况下，他们也不愿意使用新的人工智能技术。

首先，关于人工智能干预措施的质量保证和审查监督，尽管有许多法规涉及医疗保健技术（如包括电子健康记录和收入周期管理的医疗技术），但在它们如何具体解决人工智能使用（一般和特定情况）方面仍然存在清晰度和范围上的差距。例如，一些人工智能技术可能超出了现有的医疗设备监管范围。对非医疗设备的健康信息技术的监

管权部分属于 ASTP/ONC。如在“医疗产品开发、安全和有效性”章节中所述，ASTP 的 HTI - 1 最终规则本身并不创建审批流程，但它确实制定了政策，要求某些认证健康 IT (如电子健康记录) 对其提供的基于人工智能的技术保持透明。从 2025 年 1 月 1 日起，最终规则中确定的法规要求，对于作为健康 IT 产品一部分提供的、经认证符合 45 CFR 170.315 (b)(11) 的任何决策支持干预技术 (包括基于人工智能的决策支持干预)，必须提供特定的“源属性”信息。越来越多的健康 IT 中的人工智能工具可能超出当前监管范围，包括某些集成到电子健康记录中的人工智能决策支持工具 (如预约未到预测算法) 和健康计划和保险发行商为利用管理和预先授权部署的人工智能算法。这些不符合 FDA 监管的“设备”法定定义的工具目前可能不接受监管审查、验证或测试。

此外，HTI - 1 最终规则适用于基于人工智能的技术，无论设备定义、用例 (如临床、行政) 或风险类别如何。HHS 旨在进一步完善其涵盖人工智能技术的监管框架，以促进安全和可信的使用。

此外，关于责任，虽然在医疗实践中使用技术的责任方面有相当多的经验，但人工智能 (特别是生成式人工智能如大型语言模型) “引发了与旧形式的临床决策支持或在线研究医疗问题的方式不同的独特问题”。在人工智能中使用患者数据引起了医疗专业人员和患者的担忧。这些包括它如何在模型开发中使用、患者对提供者和开发者的数据存储的同意，以及患者何时被告知使用情况。模型输入和训练的持续更新使得难以建立事实模式和 / 或重现证据规则所需的特定事件或场景。

关于患者数据使用，当涵盖实体或业务合作伙伴为人工智能开发或维护使用或披露受保护健康信息 (PHI) 时，必须遵守 HIPAA 隐私和安全规则。根据 HIPAA，使用和披露 PHI 需要患者书面授权，除非允许用于某些特定目的，如治疗、支付或医疗保健操作。当 PHI 用于涉及人工智能的研究时，根据披露的 PHI 类型和进行的研究类型，HIPAA 隐私规则可能要求个人授权使用或披露 PHI，或由机构审查委员会 (IRB) 提供授权的豁免或变更。与人工智能开发者共享 PHI 也可能会带来额外的复杂性。最终，为人工智能模型使用或披露患者数据 (包括 PHI) 需要根据具体情况进行评估和管理，以确保符合 HIPAA 和其他隐私法规。“促进可信人工智能开发和道德及负责任使用”部分进一步讨论了患者安全和隐私。

最终，补充指南和法规，同时提高 HHS 部门在监督和质量保证方面的清晰度，将增强在交付和融资中使用安全和适当的人工智能用例的信心。

HHS 近期优先事项：

- 提供关于医疗保健中人工智能使用应如何遵守隐私和安全标准 (包括 HIPAA) 的额外指南。这将包括提供关于在 HIPAA 背景下重新识别风险的指南，并明确用于人工智能的数据何时需要患者授权 (即研究)。为执行此优先事项，HHS 将与其他联邦机构合作，创建人工智能应用中隐私和安全的统一标准和框架。
- 在适用的现有 HHS 和部门权限内，提供关于使用人工智能的临床医生和医疗服务提供者的责任考虑的额外指南。

- 提供关于临床医生和支持人员在与人工智能互动中应具有适当方法和角色的指南和框架（例如，根据人工智能应用的风险级别确定与技术的角色适配性）。
- 继续澄清并提高利益相关者对适用的监督和监管结构的认识。

3. 确保人工智能开发者和潜在部署者清楚覆盖和支付确定流程，以鼓励人工智能开发

背景：

由于许多提供者已经面临经济压力，他们对投资或使用新的和新兴信息技术的兴趣有限，特别是在不能保证服务支付的情况下。明确的人工智能支持服务支付框架将影响人工智能在医学中的更广泛使用，因为提供者可能会在财务上更有动力利用这些技术。

提高人工智能服务支付框架的清晰度将需要政策制定者向技术开发者、设备制造商、临床医生和患者传播信息。明确支付确定流程可以支持众多优先事项，包括为获取创新技术提供信息、减少开发者和制造商的不确定性、保护联邦计划受益人的安全、管理联邦资金以及在存在差距的地方鼓励证据开发。

HHS 行动迄今为止（非详尽）：

- 截至 2024 年 5 月，CMS 通过现行程序术语（CPT®）和新技术附加支付（NTAP）在医疗保险住院预付费系统（IPPS）下为至少八种人工智能 / 机器学习支持的设备建立了支付途径，这不到 FDA 授权的基于人工智能产品的 5%。CMS 已采取措施确保医疗保险覆盖确定和支付途径对于创新（包括人工智能支持的创新）是清晰的。
- 如在“医疗产品开发、安全性和有效性”章节中所述，CMS 对软件即服务的支付为医院门诊部门通过附加代码建立了支付途径。
- CMS 的新兴技术过渡性覆盖（TCET）（CMS - 3421 - FN）在 2024 年 8 月最终确定了 TCET 途径，以促进医疗保险受益人更安全和更可预测地获取新技术，并进一步减少覆盖的不确定性。

HHS 近期优先事项：

- 召集关键利益相关者，告知联邦保险计划的覆盖流程和要求（例如政策制定者、技术开发者、设备制造商、临床医生和患者）。
- 提供关于向联邦受益人提供的新人工智能产品和服务的覆盖确定流程的指南和清晰度。
- 为人工智能开发者制定关于支付和覆盖决策的证据标准的指南。

3.6.2 促进可信人工智能开发和道德及负责任使用

医疗保健交付中人工智能的主要重点是确保患者安全、安保和隐私。AHRQ 将患者安全定义为旨在保护护理管理中患者的多方面学科。必须彻底减轻与人工智能相关的潜在患者不良事件（无论是由人工智能支持的工具执行的错误操作还是医疗工作人员错误使用人工智能支持的工具导致）。在医疗保健交付中，一些增加人工智能可信度和安全性的

方法包括确保在人工智能决策过程中有人类参与，确保模型及其由提供者和支付者的使用是透明、可解释和可说明的，并为其使用建立明确的规则。人工智能系统缺乏可解释性可能会导致临床医生的怀疑、过度依赖或拒绝。

以下是 HHS 可以支持人工智能安全使用的优先领域：

1. 加强执法并澄清与现有要求相关的指南
2. 提供与内部治理相关的指南和支持
3. 促进外部评估、监测和透明度报告
4. 增强基础设施以确保患者安全

以下，HHS 讨论了背景、其迄今为止的行动以及促进医疗保健交付中可信人工智能开发和道德及负责任使用的计划。

1. 加强执法并澄清与现有要求相关的指南

背景：

如前文所述，HHS 及其部门可以通过澄清现有医疗保健法规并积极执行某些人工智能应用可能违反的现有法律要求，在人工智能背景下促进使用并保护受益人。在存在现有联邦要求是否得到适当应用的问题的情况下，这些努力将提供更安全、更好的患者体验。

鉴于人工智能风险的迅速扩大，仅对现有法规进行澄清可能并不总是足够的。HHS 可以制定新的杠杆、规则和程序，以确保医疗保健组织和人工智能开发者在人工智能生命周期的每个阶段（包括设计、开发、部署、维护和退役）都遵循最佳实践风险缓解原则。

HHS 迄今为止的行动（非详尽列举）：

- 美国医疗保健研究与质量局（AHRQ）的人工智能针对第 14110 号行政命令，并作为患者安全组织计划的一部分，开发了一个医疗保健安全项目（关于患者安全的额外专门讨论见下文小节），以便快速开发以人工智能患者安全为重点的数据、分析和资源。该项目有助于集体追踪和识别在医疗保健环境中部署的人工智能可能导致不良事件的情况，并提供了一种在未来从这些事件中学习的方法。2005 年的《患者安全和质量改进法案》（美国法典第 42 卷第 299b - 21 条及后续条文）设立了患者安全组织计划，也为组织提供了一定的法律保护，使其能够分享患者安全事件的信息以改善护理，而不必担心这些信息会在法律或行政程序等情况下被用来对其不利。508
- HHS 关于州、地方和部落政府（STLT）在公共福利管理中促进自动和算法系统中人工智能负责任使用的计划包括一些建议，如进行影响评估以确定人工智能的预计效益和风险、衡量系统训练、测试和预测中使用的数据的质量和适当性，以及咨询工作人员并为所有员工提供关于开发、使用、增强和维护自动和算法系统的充分培训。509

- 《健康计划和活动中的非歧视最终规则》(《患者保护与平价医疗法案》第 1557 节 [“第 1557 节”]) 禁止在某些健康计划和活动中存在歧视, 并且像其他联邦民权法一样, 第 1557 节适用于人工智能、临床算法、预测分析和其他工具的使用。

2 提供与人工智能本地治理相关的指南和支持

背景: 部分医疗服务和融资组织已建立了自己的治理框架和审查机制, 但在其他情况下, 风险评估流程和清晰的治理结构可能缺失或不够严格, 无法充分保护患者权益。

HHS 已采取行动 (非详尽列举):

- HHS 和各部门发布了如《值得信赖的人工智能手册》和《CMS 人工智能手册》等, 为组织安全实施人工智能开发提供具体指导。
- 医疗保健研究与质量局 (AHRQ) 发布了《解决算法偏见对健康和医疗保健中种族和民族差异影响的指导原则》以及《数字医疗公平框架和实施实用指南》, 帮助组织在人工智能生命周期的各个阶段减少偏见, 促进公平。

HHS 近期优先事项:

- 在 HHS 职权范围内, 支持制定针对医疗组织的风险管理和内部人工智能治理指南, 以现有政策、治理和风险管理框架为基础。
- 探索利用联邦项目和激励措施, 要求或鼓励医疗服务和融资组织建立内部治理机制和评估实践。
- 制定评估人工智能工具风险的最低标准, 包括基于设备潜在影响的风险分层指南和相应的监测节奏与指标。
- 制定医院指南和资源, 以识别、管理和缓解与人工智能相关的安全、偏见或有效性问题。

HHS 长期优先事项:

- 通过出版物、研究和标准, 持续向公众和临床团队宣传可信赖和安全的人工智能, 包括如何解读人工智能、向患者沟通干预措施、识别可能出现的不良事件以及如何通过现有系统报告这些事件。
- 召集并 / 或支持与行业专家就人工智能风险和适当的风险管理方法举行公开会议和对话。

3 促进外部评估、监测和透明度报告, 以提高人工智能的质量保证

背景: 在现实医疗环境中测试和评估人工智能的效果具有挑战性, 这是由于人工智能在不同临床领域的快速扩展以及临床医疗数据的常见问题 (如某些疾病患病率低、缺乏或难以获取真实数据) 所致。此外, 人工智能的持续学习和适应能力也给监管带来了独特挑战。

HHS 已采取行动（非详尽列举）：

- 《健康技术创新 - 1 最终规则》(HTI - 1 Final Rule) 为电子健康记录 (EHR) 开发者及其在 ONC 健康信息技术认证计划下认证的系统中的某些人工智能建立了一套透明度要求。

其他行业已采取行动（非详尽列举）：

- 多个组织合作开展倡议，召集医疗服务利益相关者应对挑战，并发起与人工智能相关的倡议。
- 值得信赖和负责任的人工智能网络 (TRAIN) 和健康人工智能联盟 (CHAI) 等组织致力于推动人工智能的合理使用和发展，并制定了相关模板和标准。

HHS 近期优先事项：

- 与行业合作，明确 HTI - 1 决策支持干预的源属性（即“模型卡”）要求的标准化指标、信息和数据的共识方法，以增强透明度要求。
- 考虑通过激励措施，鼓励披露医疗服务提供者和支付者使用人工智能及其对医疗服务获取或质量的影响的性能信息，扩大人工智能性能信息的可获取性。
- 探索利用联邦项目和激励措施，鼓励部署人工智能的支付者和提供者组织建立外部问责机制，包括促使人工智能部署者接受由无利益冲突的认证实体进行的独立外部算法审计，激励开发者和部署者提高性能透明度，以及与私营部门现有的和新兴的验证、监测和透明度工作进行合作。

4 加强基础设施建设，确保患者安全、保障和隐私

背景：随着人工智能在医疗服务中的应用，维护患者的安全、保障和隐私成为关键但复杂的挑战。患者对影响其医疗决策的透明度需求日益增加，同时人工智能也带来了新的安全和隐私漏洞。

HHS 已采取行动（非详尽列举）：

- AHRQ 的医疗保健安全计划 (AI in Healthcare Safety Program) 采取措施分析和汇总各类人工智能事件的数据，并鼓励更多组织与患者安全组织 (PSOs) 合作。
- AHRQ 通过患者安全数据库网络 (NPSD) 对患者安全事件进行探索性分析，以了解人工智能在医疗环境中可能产生的意外影响。

HHS 近期优先事项：

- 扩展 PSOs 的能力，通过教育、资源共享和开发，帮助医疗服务提供者从可能与人工智能相关的不良影响中吸取教训并加以预防。
- 探索鼓励在医疗保健安全计划中提交可能与人工智能相关事件数据的机制。

HHS 长期优先事项：

- 将 NPSD 用作医疗环境中因人工智能部署导致的患者安全事件的“中央跟踪存储库”。
- 考虑扩展 AHRQ 的医疗保健安全计划，以维持和推进对 NPSD 进行分析的初始项目。
- 促进 AHRQ 关于减轻算法中种族偏见的研究。
- 考虑扩大赠款项目和资助机会公告（NOFOs）。
- 继续评估人工智能对医患互动（如直接面对面时间、患者信息收集）的潜在影响。

3.6.3 普及人工智能技术和资源

为实现人工智能加速医疗服务获取和公平性的目标，技术及其实施的理解必须易于获取。如果在训练数据中未充分考虑某些患者群体代表性不足所导致的偏见，服务不足的地区可能无法从人工智能中获得充分益处。HHS 可在以下领域采取行动，特别关注可能已受数字鸿沟影响的利益相关者群体：

1. 通过技术支持和与为服务不足人群提供服务的医疗服务组织合作，促进公平获取

- **背景：**人工智能对服务不足人群的潜在影响程度仍很大程度上未知，服务不足社区可能面临复杂且研究不足的细微问题。服务于这些地区的医疗服务提供者可能从人工智能中获益最大，但也面临缺乏专业知识和资金来部署新技术的最大障碍。
- **HHS 已采取行动（非详尽列举）：**美国国立卫生研究院（NIH）的 AIM - AHEAD 计划通过合作、研究、基础设施建设和数据科学培训，为代表性不足的社区提供人工智能访问权限，以增加人工智能研究人员和数据的多样性。
- **HHS 近期优先事项：**通过赠款或合作协议建立区域技术援助中心，为资源不足的医疗服务机构提供人工智能应用方面的帮助；传播人工智能影响评估模板、实施工具包和技术援助资源，以帮助考虑使用人工智能的医疗服务组织；资助研究，以开发在资源不足环境中使用人工智能应用的最佳实践见解。
- **HHS 长期优先事项：**召集医疗服务领域的实践社区，促进人工智能应用方面的信息共享，特别是在服务不足人群中；继续评估人工智能对医患互动的潜在影响。

2. 为医疗服务组织提供支持，解决核心基础设施和部署挑战，提高人工智能就绪程度

- **背景：**组织需要具备包括数据系统在内的先决能力和基础设施来利用人

工智能。医疗服务是一个极其复杂的系统，不同专业和利益相关者面临不同情况，基础工具工具和人工智能难以通用。此外，数据质量、数据集可用性以及技术基础设施的可及性和可负担性等问题限制了医疗组织大规模应用人工智能的能力。

- **HHS 已采取行动（非详尽列举）：**卫生资源与服务管理局（HRSA）的统一数据系统现代化倡议更新和改进数据集及数据提交、收集和分析技术；印第安卫生服务局（IHS）的数据现代化倡议与部落和城市领导人合作，更新 EHR 标准；儿童与家庭管理局（ACF）的数据战略增强了人工智能能力，并通过提高数据互操作性支持对儿童、家庭和服务不足人群的护理服务分析。
- **HHS 近期优先事项：**与行业合作，推广供利益相关者利用的开源人工智能规范；通过赠款或合作协议建立区域技术援助中心，支持资源较少的医疗服务机构进行数据和技术现代化，以提高人工智能就绪程度；传播人工智能就绪程度评估模板，帮助考虑开发人工智能解决方案的医疗服务提供者评估数据、系统和技术基础设施方面的差距。
- **HHS 长期优先事项：**更新内部数据基础设施，确保为服务不足社区提供足够且可用的信息，以制定支持策略；提供开放的去标识化数据资产，包括行政、临床、质量 / 结果和安全数据，以支持人工智能的开发、测试和验证。

3.6.4 培养人工智能赋能的劳动力和组织文化

具备人工智能知识的劳动力将有助于加速创新、管理与新工具相关的部署风险、建立适当的组织治理结构、评估训练数据中的潜在偏见、监测模型漂移、减轻不良影响，并与患者、家属和医疗服务提供者就该技术的使用进行沟通。一支有效的健康人工智能劳动力队伍需要跨职能团队，包括临床医生、生物统计学家、隐私 / 信息安全官员、分析师、采购人员和 IT 专业人员。确保个人和组织做好充分准备使用人工智能对于其安全、有效和广泛应用至关重要。

背景：目前，人工智能在大多数医疗教育、认证和继续教育项目的课程中所占比例有限。此外，大多数人工智能专业知识集中在技术组织和 / 或研究机构（如大学和大型技术组织）中。HHS 将采取措施提高医疗专业人员的人工智能知识和专业技能，确保医疗服务组织具备基础专业知识，降低实施新工具的成本，并确保其正确应用。

HHS 已采取行动（非详尽列举）：

- 《CMS 人工智能手册》包含了定义人工智能、医疗服务中的用例和趋势的教育材料，以及 CMS 正在考虑在其运营中使用的应用程序及其对患者护理的潜在影响。

- AHRQ 的内部研究项目为卫生服务研究提供了博士前和博士后教育、研究基础设施和职业发展资助及机会,并与联邦和学术合作伙伴合作开发创新课程和教育模式。

HHS 长期优先事项:

- 通过长期公共教育倡议,提高在医疗服务环境中工作的专业人员(临床和非临床)的人工智能素养。
- 在医疗会议上定期举办人工智能会议,包括由行业专家主持的研讨会、学习课程、实践工作坊和录制资源,以促进合作、学习和创新。
- 制定针对不同医疗工作者群体(如继续医学教育、学位课程)的人工智能概念培训课程和节奏的指导方针。
- 直接资助劳动力培训项目,培训现有的健康人工智能劳动力并教育下一代医疗专业人员。
- 在公共网站上分享人工智能内部培训资源,供行业内的健康人工智能专业人员在不同医疗环境中直接使用或改编。
- 继续评估人工智能对医疗劳动力的潜在影响。

3.7 结论

通过本战略规划中的行动,HHS 将帮助医疗服务组织利用人工智能扩大医疗服务获取途径并改变患者护理方式。鉴于人工智能的快速发展,HHS 将持续审查本规划的行动,并努力为医疗服务生态系统中的利益相关者提供支持。

4 公共服务提供

4.1 引言与背景

人工智能为改善美国公共服务项目的质量、可及性、互操作性、协调性和整体影响提供了机遇。人口老龄化和多样化、复杂且分散的公共福利系统以及持续的劳动力短缺凸显了人工智能在该领域的潜力。然而，人工智能在公共服务中的应用仍处于初期阶段，面临着诸多关键挑战，包括缺乏资金、过时的信息技术和数据基础设施，以及对技术对公共服务项目参与者影响的担忧。

尽管存在这些挑战，对人工智能的兴趣依然存在，83% 的政府领导人认为该技术在支持公共服务劳动力方面将变得更加重要。HHS 已发布《促进州、地方、部落和地区政府在公共福利管理中负责任地使用人工智能和自动化算法系统的计划》，但仍有进一步行动的空间。HHS 致力于在保护美国人安全的同时，最大限度地发挥人工智能的潜力，确保该技术得到负责任的测试、部署和监测。在本章中，HHS 概述了公共服务提供中人工智能的相关范围和利益相关者，探讨了人工智能在该领域的机遇和趋势，列举了潜在用例和风险，并提出了实现 HHS 部门人工智能目标的方法。

4.1.1 公共服务提供人工智能价值链的范围

促进健康与福祉	协助有复杂需求的人群	支持家庭和儿童	加强社区和经济发展
获得医疗服务、预防保健、疾病治疗、心理健康支持	为经历经济困难、无家可归、药物滥用的个人提供援助	儿童保育、早期教育、家庭支持、儿童福利服务、寄养、收养	改善当地基础设施和服务的社区倡议、经济援助

4.1.2 行动计划摘要

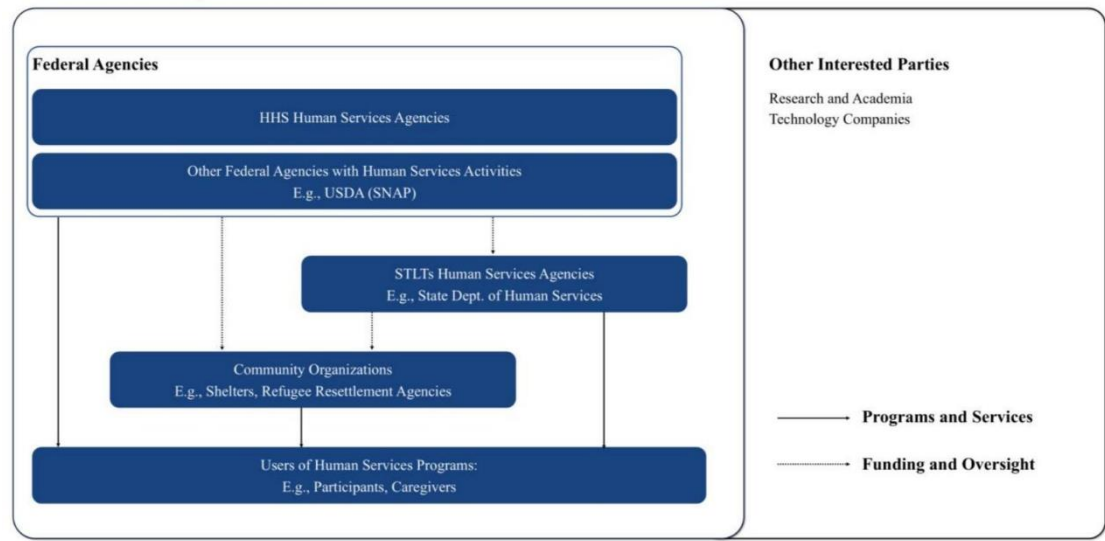
关键目标	拟议行动主题（非详尽，详见 4.6 行动计划）
1. 催化健康人工智能创新与应用	释放人工智能应用资源，实现信息技术和技术基础设施现代化；确保人工智能应用的数据质量和可用性
2. 促进可信赖的人工智能开发以及合乎道德和负责任的使用	为服务人群提供指导，平衡人工智能应用的风险与机遇，建立参与者信任

关键目标	拟议行动主题（非详尽，详见 4.6 行动计划）
3. 普及人工智能技术和资源	提高民众数字素养和数字普及率；确定跨部门合作领域，提高人工智能相关的规模经济效益
4. 培养人工智能赋能的劳动力和组织文化	提高公众服务员工的数字素养、人才水平和对使用新技术的开放态度；利用人工智能缓解公众服务劳动力短缺问题

4.2 参与公众服务提供人工智能价值链的利益相关者

美国的公众服务项目使最弱势群体、他们的照顾者和监护人受益。各类联邦、州、地方、部落（STLT）和社区利益相关者为实现这一目标做出贡献。联邦机构为 STLT 公众服务机构和社区组织提供资金以实施项目，同时自身也开展项目；STLT 为社区组织提供资金并直接提供服务；社区组织（CBOs）则利用联邦、STLT 和慈善资金实施项目。图表 10 展示了利益相关者之间示例流程的非详尽示意图，以及参与公众服务的利益相关者项目列表。请注意，该图表和列表均未涵盖所有角色和互动。有关监管指南和权限的更多详细信息，请参考其他美国卫生与公众服务部（HHS）文件。该图表反映了示例角色和关系，但角色可能因公众服务项目的不同而有所差异。

图表 10：公众服务提供利益相关者参与地图



- HHS 机构
 - ACF: 为家庭和儿童提供服务，促进儿童、家庭和社区的经济和社会福祉。

- **ACL**：支持针对有复杂需求人群的项目，特别是老年人和残疾人。
- **CMS**：管理联邦医疗保险计划（如医疗保险和医疗补助），制定相关项目的参与条件，并可为特定设备或服务提供报销。
- **SAMHSA**：致力于促进健康与福祉，包括与自杀预防、心理健康以及药物滥用治疗和预防相关的服务。
- **HRSA**：为服务不足人群提供基本医疗服务，重点关注促进健康和协助有复杂需求人群的服务。
- **IHS**：为美国印第安人和阿拉斯加原住民提供全面的医疗服务体系，确保提供符合文化背景的公共卫生和公众服务。
- **其他联邦机构**：HHS 与许多其他联邦部门密切合作，如农业部和住房与城市发展部。
- **STLT 政府公众服务机构**：负责管理项目并提供公共福利，通常管理联邦项目如补充营养援助计划（SNAP）和妇女、婴儿和儿童特别补充营养计划（WIC）以及 HHS 项目。
- **CBOs，包括社区行动机构**：直接向公众提供公众服务项目和福利。
- **参与者及其照顾者和监护人**：2023 年，估计有 9910 万人（占美国人口的 30%）使用了各类公众服务项目。
- **技术公司**：包括专注于人工智能基础设施（如云存储）的公司、大型多元化科技公司、数字解决方案供应商和白帽黑客，为利益相关者使用人工智能提供基础设施和服务。
- **研究机构**：通常与联邦机构、STLT 或 CBOs 合作，进行试验和评估，以了解公众服务干预措施的证据，并设计和测试潜在项目。

4.3 人工智能在公众服务提供中的应用机会

人工智能在公众服务中可改善服务体验和质量、提高资金分配的速度和质量、增强公众服务劳动力的能力、增加服务的可及性，并提高互操作性以改善服务协调。这些机会由多种因素驱动，包括人口结构变化、复杂的公共福利生态系统和劳动力短缺。

1. **改善服务体验和质量**：符合条件的参与者在获取公众服务项目时面临挑战，且满意度较低。人工智能可通过协助参与者匹配项目、加快申请流程、提高福利发放速度和增强参与者支持体验来解决这些问题并提高满意度。
2. **提高资金分配的速度和质量**：数十亿美元通过 HHS 流向 STLTs、社区组织和美国民众。通常，这些资金分配得越快，公共福利和重要服务就能越快提供。正如

COVID - 19 疫情期间推出的援助计划所示，快速有效地提供资金有可能拯救生命和生计。人工智能有潜力提高 HHS 向其他利益相关者分配资金的速度和准确性，确保资源分配公平且与最需要的地区相关联。

3. **增强公众服务劳动力的能力：**公众服务部门在招聘和留住关键劳动力方面面临挑战，且需求不断增长。例如，美国劳工统计局预计未来十年美国每年将短缺 67300 名社会工作者，其他估计显示缺口更接近 100000 名。劳动力短缺可能导致等待时间延长和服务减少。与此同时，美国人口老龄化，预计将有更多人使用公众服务项目，这可能给劳动力带来压力，增加对服务的需求，并更加强调高效的福利提供。人工智能可以通过自动化常规任务、使用自然语言处理（NLP）处理叙述性信息（如客户笔记、会议记录、访谈记录）以及起草文件等方式，辅助公众服务劳动力的工作流程。在客户支持中心这一类似场景中，美国国家经济研究局的一项研究发现，使用基于人工智能的对话助手可将工人生产率提高 14%。在公众服务领域实现类似的生产率提升，即便存在员工短缺的情况，也能让工作人员有更多时间投入到增值任务和与参与者的互动中，从而提高工作效率。HHS 也认识到与潜在员工岗位替代相关的担忧，并在下文列出了监测劳动力和服务影响的行动。
4. **增加服务的可及性：**越来越多不同背景的人有资格享受公众服务，但许多人在获取这些项目时遇到困难。据城市研究所称，十分之四的成年人报告在申请公众服务（包括贫困家庭临时援助计划（TANF）和 SNAP）时遇到困难。多种因素可能导致申请的可及性挑战，例如福利申请需要较高的词汇量、健康素养或金融素养。然而，根据美国社区调查，2600 万美国居民（约占人口的 9%）英语水平有限。项目获取方面的挑战可能使符合条件的参与者在需要时无法获得服务。例如，在 COVID - 19 疫情期间，尽管 WIC 的资格人数增加，但 2020 年至 2021 年期间参与人数仅增长了 2%。人工智能可以通过增加服务的可及性并帮助实现公平目标，为公众服务提供生态系统中的利益相关者提供帮助。人工智能应用通过视觉辅助和字幕等技术提高了其他领域的可及性。此外，生成式人工智能的进步提高了自动语言翻译的准确性和文化适应性。公众服务工作人员可能无法完全依赖这些工具，但他们可以根据参与者的需求调整翻译模型，以覆盖因语言障碍而长期得不到充分服务的社区。
5. **提高互操作性以改善服务协调：**大量与公众服务相关的数据以叙述形式（如案例笔记）收集或手动转录（如在避难所中）。这些记录不易搜索，需要人工审查，从而影响了准确需求评估、服务提供、系统分析或机构间互操作性的数据质量。互操作性的另一个挑战是美国公共福利系统的复杂性和多面性。希望申请项目的人必须遵守各种行政和项目要求，但这些要求在各州和系统之间存在差异。人工智能技术的进步，包括光学字符识别、自然语言处理和大语言模型，有可能将数据转换为结构化格式，从而更轻松地改善服务提供并在机构间共享数据。此外，更高质量的数据可以使诸如综合福利系统之类的人工智能应用将服务提供转变

为以个人为中心的设计，从而将医疗保健、住房、家庭援助和食品安全等方面的福利进行协调并共同提供。

4.4 公众服务提供中人工智能的应用趋势

当前趋势表明，公众服务领域的人工智能应用尚处于起步阶段，但州、地方和社区组织对试点创新技术的兴趣日益浓厚：

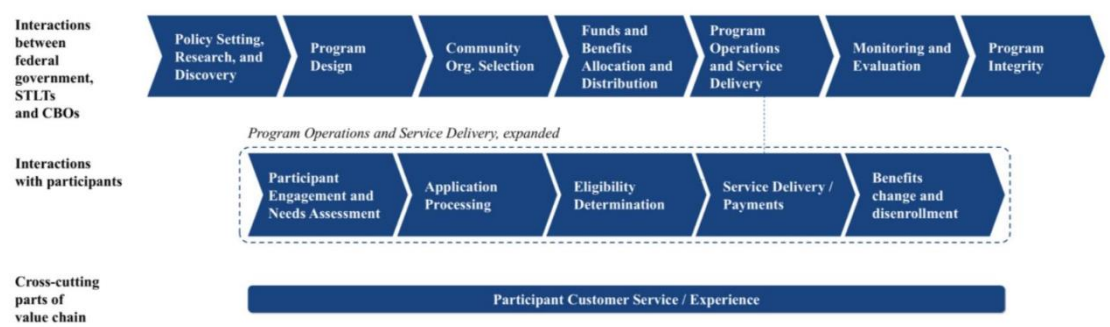
1. **州、地方和社区组织对使用人工智能感兴趣，但仍处于早期阶段，重点在于构思和合作：**多个非营利组织已建立人工智能实践，表明该领域普遍存在热情。以下列举了公众服务生态系统中的一些行动示例，但并非详尽无遗：
 - 美国数字响应组织推出了工具，帮助州和地方政府安全使用生成式人工智能，以便更好、更快地完成工作。
 - 公共部门人工智能中心推出了“滚动提示”项目，使从事人工智能工作的公司和组织能够与州卫生和公众服务部门的领导分享如何在该领域应用其技术的想法。
 - GovAI 联盟制定了政策模板和其他资源，以支持州、地方政府实施治理，从而负责任地试验和使用人工智能。该联盟代表了五百多个主要来自城市和地方政府的机构。
2. **尽管人工智能提供了诸多机会，但其在公众服务领域的使用率仍然较低：**与其他领域相比，人工智能在公众服务生态系统中的大规模使用（超出试点阶段）程度较低。虽然一些私营部门和非营利组织推出了相关项目，但这些项目通常范围有限，或针对特定地区或人群。例如，一些利用生成式人工智能辅助福利申请的平台，但它们往往范围有限（例如仅关注带薪休假政策），且未整合到各州的申请流程中。其他公众服务项目的实施示例包括社交或辅助机器人以及基于生成式人工智能的面试模拟。
3. **使用率低部分归因于对公益项目或其他非营利合作的依赖：**公共部门服务提供机构在进行试点时依赖外部支持，这些试点通常针对风险较低的用例，以减轻行政负担或支持研究。例如，伊利诺伊州就业安全部与美国数字响应组织合作，使用基于机器学习的翻译软件改进失业保险政策的翻译。此外，HHS 在公众服务领域看到了更多与人工智能相关的研究活动，这些活动受到激励，旨在开发创新方法，例如在儿童福利领域预防儿童受到虐待。
4. **对人工智能潜在负面影响的担忧可能限制其在公众服务中的使用：**公众服务领域的利益相关者在没有进行风险评估和制定严格要求以应对潜在不利影响之前，不愿使用人工智能工具。这些都是重要的保障措施，因为数据驱动的偏见可能会进一步加剧现有的不平等，使服务对象面临结果恶化和进一步被排斥的风险。此外，

正如《公共福利中负责任使用人工智能计划》所规定的，人工智能应用存在影响权利和安全的风险，例如自动拒绝项目申请。为应对这些风险，利益相关者可能会对技术供应商和服务合作伙伴提出更高要求，尽管这很重要，但与其他领域相比，这可能导致人工智能在公众服务领域的使用滞后。因此需要权衡利弊。

4.5 人工智能在公众服务提供中的潜在用例和风险

公众服务提供中的价值链因项目和组织而异。例如，离家出走和无家可归青少年项目、“领先开端”项目、难民安置项目和儿童福利服务所涉及的活动类型和顺序各不相同。以下是公众服务核心功能的概述。

图表 11：公众服务价值提供链



环节	具体说明
公 众 服 务 交 付 价值链	联邦政府、州、地方和部落政府（STLTs）与社区组织之间的互动涵盖政策制定、研究与发现、项目设计、社区组织选择、资金与福利分配、项目运营与服务交付、监测与评估、项目完整性等功能组件。与参与者的互动包括参与者需求评估、参与和申请处理、资格确定、支付、服务交付以及福利变更和退出。此外，还有跨职能的客户服务 / 体验环节。
项 目 运 营 和 服 务 交 付 (扩展)	包括与参与者互动的各个功能组件，从需求评估到福利变更和退出的整个过程。

4.5.1 公众服务交付价值链上的人工智能用例

在下面的表格中，HHS 列出了公众服务交付价值链中人工智能的潜在益处和风险，但并非详尽无遗。请注意，以下详细介绍的用例突出了该领域各种利益相关者现有或潜在的

人工智能使用方式。有关 HHS 及其部门如何使用人工智能的详细信息，请参考《HHS 人工智能用例清单 2024》。HHS 指出，人工智能只是公众服务交付利益相关者可用的多种技术工具之一，过度依赖人工智能可能会带来需要充分解决的风险。此外，许多技术（如大语言模型）仍在公众服务环境中进行潜在风险评估。HHS 在其他文件中，包括《HHS 值得信赖的人工智能手册》和《公共福利中负责任使用人工智能计划》，已经讨论了使用人工智能的风险考虑因素。除了潜在用例，HHS 还列出了需要考虑的潜在风险，同样并非详尽无遗。HHS 将在本章后面提出的行动中考虑缓解已识别风险的措施。

1. 联邦政府、STLTs 和社区组织之间的互动

○ 功能组件 1：政策制定、研究与发现

- **潜在用例（非详尽）**：能够综合多个不同数据集以辅助政策评估和制定的工具；从项目测量数据、人口统计数据和其他领域生成人工智能驱动的见解，为政策制定提供信息，这类工具可以综合多种不同数据集，为政策评估和新政策制定提供信息。
- **潜在风险（非详尽）**：存在第三方风险，例如通过第三方供应商导致项目数据泄露；可能发生敏感人口数据的数据泄露；第三方数据存储可能存在访问风险。

○

○ 功能组件 2：项目设计

- **潜在用例（非详尽）**：用于辅助项目设计的人工智能驱动的测量和数据分析工具；利用政策制定阶段的人工智能驱动见解，为项目设计和交付提供最佳实践指导；人工智能可支持资源映射、人口统计和需求的地理空间分析以及预测建模，为项目设计和组织选择提供信息。
- **潜在风险（非详尽）**：存在可解释性和问责制风险，例如基于黑盒算法分配项目资源；人工智能应用可能存在输入缺陷，且缺乏适当的保障措施，导致用户无法理解决策过程或训练数据，从而为潜在参与者分配不匹配的资源。

○ 功能组件 3：社区组织选择

- **潜在用例（非详尽）**：用于管理社区组织网络和识别新合作伙伴的评估工具，例如将个案工作者在网络监测期间记录的叙述性和语音笔记转换为可更高效评估和长期分析的数字数据；通过网络抓取获取社区组织基准数据；利用人工智能驱动的网络搜索进行快速费率基准测试、服务可用性搜索以及组织过往历史查询，为网络选择、赠款审批和谈判提供信息。

- **潜在风险（非详尽）：**由于对非结构化数据的错误解释，可能导致误判，例如对网络抓取数据的不准确结构化导致提案评估错误；人工智能驱动的网络抓取可能会不准确地评估社区组织网站和公共参考资料中的数据，从而得出错误或误导性结论，影响赠款授予或合作伙伴评估。
- **功能组件 4：资金与福利分配**
 - **潜在用例（非详尽）：**在项目允许的资金灵活性范围内，评估赠款申请并将资源分配到需求最高地区的工具；能够更快、更明智地阅读自由裁量赠款申请的工具；预测资金短缺的分析工具；利用人工智能驱动的评估，确定项目和组织在资金利用上的不足或过度情况，以改善公众服务生态系统中的支出分配。
 - **潜在风险（非详尽）：**可能因输出错误导致资金分配不公平，例如基于有缺陷的人工智能算法进行资金分配，导致 STLTs 和 CBOs 资源短缺；人工智能算法基于错误的评估进行资金分配，可能导致 STLTs 和 CBOs 获得的资源不足，无法开展项目和分配福利。
- **功能组件 5：项目运营与服务交付：**此部分活动从参与者参与和需求评估到福利变更和注册，具体用例和风险在“与参与者的互动”功能组件中详细说明。
- **功能组件 6：监测与评估**
 - **潜在用例（非详尽）：**利用非结构化数据评估项目有效性的工具，例如将项目实施过程中记录的叙述性笔记（如无家可归者收容所的纸质记录、行为健康咨询中的案例笔记）转换为可用于测量和评估的数字记录；基于项目数据、个案工作者笔记和外部来源的人工智能驱动数据分析，以衡量项目产出和成果，提供超越描述性或领先指标的见解。
 - **潜在风险（非详尽）：**可能出现误判或错误输出，例如人工智能生成的笔记中的虚构内容导致个案工作者评估错误；转录个案工作者项目笔记时的人工智能虚构内容可能导致数据错误分类或项目绩效的错误测量，进而影响政策决策。
- **功能组件 7：项目完整性**
 - **潜在用例（非详尽）：**对公众服务项目组合进行数据驱动的持续监测，以发现异常情况，例如检测福利使用、承包商行为中的异常模式，以及潜在的欺诈、浪费和滥用行为，使用人工智能驱动的分析替代手动调查；具备评估项目完整性并标记整个供应商网络

中潜在异常活动能力的驱动仪表盘。

- **潜在风险（非详尽）：**可能存在人工智能模型使用不当和输出错误的情况，例如不适当的人工智能欺诈检测应用导致错误的项目调查；为其他目的（如金融服务中的欺诈检测、支付处理或验证）开发的人工智能应用，在公众服务（如项目欺诈检测）中使用时可能导致错误的欺诈调查。

2. 与参与者的互动

○ 功能组件 8：参与者参与和需求评估

- **潜在用例（非详尽）：**协助公众服务劳动力更好地预测人口需求并与参与者沟通的工具，例如预测高风险个人和人群，并尽早联系他们进行注册、干预和提供全方位服务（如心理健康危机支持）；为个案工作者提供实时语言和跨文化翻译工具，使其能够与讲不同语言的参与者互动，并帮助识别跨文化沟通障碍。
- **潜在风险（非详尽）：**可能出现错误输出，例如不准确的实时翻译导致参与者和个案工作者之间的信息传达错误，在讨论法律文件或护理事宜时尤其可能造成严重的沟通差距。

○ 功能组件 9：申请处理

- **潜在用例（非详尽）：**能够更快速、准确地处理申请并向参与者提供通俗易懂信息的平台，例如基于有限的人口统计和社会因素，帮助人们了解可用项目并初步确定其资格的预测性资格确定工具；尽可能自动化申请任务并利用数据连接的简化申请处理工具。
- **潜在风险（非详尽）：**可能存在算法偏见，导致决策错误，例如对情况相似的人做出不准确的福利判定；如果人工智能应用于公众服务交付的各个环节（如项目资格确定、欺诈检测或风险分层），可能会对人员进行错误分类。

○ 功能组件 10：资格确定

- **潜在用例（非详尽）：**连接多个不同的公众服务系统，以改善福利选择并加快服务交付；预测个人在注册一个项目后可能符合并需要其他项目的可能性，并在注册时以通俗易懂的语言推荐这些服务；人工智能驱动的跨多个机构（如医疗保健、公众服务、住房和食品安全）的项目数据、应用、资格确定和服务交付的集成。
- **潜在风险（非详尽）：**与功能组件 8“申请处理”中概述的风险相同。

- **功能组件 11：服务交付和支付**

- **潜在用例（非详尽）：**人工智能生成的服务内容和支持平台，以扩大项目的覆盖范围和有效性，例如为老年人提供社交隔离游戏、为认知行为疗法提供数字治疗干预（如聊天机器人）、为心理健康项目提供对话代理等；在老年人或残疾护理中使用人工智能机器人，社交机器人可帮助治疗老年人的隔离或痴呆症状，辅助机器人可协助完成日常任务，包括个人卫生和移动。
- **潜在风险（非详尽）：**可能存在偏见或错误输出，例如基于有偏见的数据或有缺陷的算法，对项目参与者是否适合某个教育或劳动力培训项目做出不当评估，导致错误的安置或项目提供；人工智能驱动的案例复杂性评估不准确可能会因资源分配不当而加剧劳动力挑战。

- **功能组件 12：福利变更和退出**

- **潜在用例（非详尽）：**主动的参保人管理，以确保准确的重新注册和福利变更，例如使用人工智能确认参保人信息并协助确认资格、办理退出或重新注册手续；利用跨多个机构整合的数据预测参与者资格（如福利悬崖）何时会发生变化，并以通俗易懂的语言主动通知。
- **潜在风险（非详尽）：**可能会加剧参与者的信任担忧和对人工智能的怀疑，例如为预测福利变更（如福利损失）的人工智能模型过度收集数据，可能会增强不信任感，特别是对于对公众服务项目已经持有负面看法的代表性不足群体。

3. 价值链的跨职能部分

- **功能组件 13：客户服务 / 体验**

- **潜在用例（非详尽）：**改善公共服务人员与参与者互动并更简单准确地提供支持的客户支持工具，例如生成式人工智能驱动的聊天机器人或虚拟助手，能够以多种语言通俗易懂地回答参与者或潜在申请人的问题，并辅助基本资格预测和融入申请处理流程；利用生成式人工智能和与非结构化个案工作者笔记及呼叫中心反馈的连接，进行情感分析并识别参与者咨询和呼叫中的趋势和共同主题的综合参与者反馈工具；为社区组织和 STLTs 提供的聊天机器人或虚拟助手，帮助他们了解政策、获取指导并回答与项目相关的问题；利用人工智能驱动的分析，了解哪些服务渠道、时间和其他情况需要不同的服务能力，并优化劳动力以满足需求。

- **潜在风险（非详尽）：**可能出现错误输出，例如内部或外部面向参与者的人工智能聊天机器人提供错误信息，导致参与者获取项目的障碍，或使他们认为不符合项目资格。

4.6 行动计划

鉴于公共服务提供中人工智能格局的不断演变，HHS 已采取多项措施，包括发布州、地方和部落政府在公共福利中使用人工智能的新指南、通过公私合作伙伴关系分享实践经验以及提供赠款资金以促进负责任的人工智能应用。以下行动计划遵循支持 HHS 人工智能战略的四个目标：1. 催化健康人工智能的创新与应用；2. 促进可信赖的人工智能开发以及合乎道德和负责任的使用；3. 普及人工智能技术和资源；4. 培养人工智能赋能的劳动力和组织文化。对于每个目标，行动计划提供了背景信息、HHS 和其他相关联邦机构迄今为止的行动概述，以及 HHS 将采取的近期和长期优先事项。HHS 认识到，随着技术的发展，该行动计划需要定期修订，并致力于提供结构和灵活性，以确保产生长期影响。

4.6.1 催化人工智能创新与应用

HHS 可以通过以下领域的相关机会来促进人工智能的创新与应用：

1. 释放人工智能应用资源并实现信息技术和技术基础设施现代化

- **背景：**公共服务领域的赠款和合同通常不会为人工智能相关投资分配资金，也不将具备特定信息技术能力作为授标的条件。总体而言，该领域面临资金和劳动力短缺问题，许多利益相关者认为无法在一年内满足服务需求。州、地方和社区组织可能缺乏可用于人工智能投资或技术基础设施改善的资金。由于这种持续的资金短缺，非营利组织在信息技术基础设施上的投入少于私营部门，尽管技术更先进的非营利组织更有可能实现其使命。这种投资不足导致机构和社区组织的信息技术基础设施过时，或过度依赖模拟和纸质记录保存方式。想要将人工智能融入运营的组织可能需要跨越多个信息技术成熟阶段。州、地方和社区组织若在缺乏适当技术基础的情况下对信息技术进行转型投资，可能会面临额外挑战，例如因人工智能使用故障排除而导致服务质量下降。更多用于人工智能应用的资源将带来诸多机会，包括增强互操作性、提高资金分配的速度和质量，以及改善服务质量和体验。
- **HHS 已采取的行动（非详尽列举）：**
 - HHS 的《公共福利中人工智能的负责任使用计划》：
 - 概述了对州、地方和部落政府在公共福利中推广人工智能使用的额外支持领域，包括提供有关州、地方和部落政府

可获得的资金信息。

- 为州、地方和部落政府以及供应商在公共福利管理中有效使用人工智能推荐了具体的推动因素，这些推动因素包括改善信息技术基础设施、高质量数据和适当的保障措施。
- 探索为试图在公共福利中实施负责任人工智能的州、地方和部落政府提供技术援助，以提高其合理利用人工智能的能力，并根除和缓解相关风险。
- 利用现有合作伙伴关系并建立新的合作关系，以协调人工智能的推广和应用。HHS 与联盟、咨询委员会和其他组织建立了现有合作关系，并正在建立新的合作关系，以帮助跨辖区分享最佳实践和经验教训，包括错误案例。这种信息共享可以缩短新使用者的学习曲线，并提供实际操作指南，包括政策、治理和人工智能工具采购的模板。
- 为社区组织提供赠款资金，通过人工智能应用提高服务质量。儿童与家庭管理局（ACF）和老龄化管理局（ACL）正在资助一些组织，这些组织通过多种方式使用人工智能来改善其运营或项目交付，例如在辅助生活设施中使用机器人技术或推出人工智能驱动的聊天机器人。
- **其他联邦机构已采取的行动（非详尽列举）：**美国农业部（USDA）在 2024 年 4 月发布了《州、地方和部落政府在公共福利中使用人工智能的框架》。该框架与 HHS 的《公共福利中人工智能的负责任使用计划》类似，为州、地方和部落政府使用人工智能提供了指导方针，包括确定人工智能应用的益处和目标、与供应商的互动，以及数据的负责任使用和信息技术系统设计。
- **HHS 近期优先事项：**
 - 为资源较少的州、地方和社区组织确定在公众服务中应用人工智能（包括信息技术现代化、数据质量改进或其他投资类型的赠款计划）的可用资金机会。
 - 探索与私营部门的合作，这些合作可以为有兴趣在公众服务项目中使用人工智能应用和实现信息技术现代化的 HHS、州、地方和社区组织提供技术援助。
 - 汇编并提供信息技术现代化的最佳实践实施方案，包括《公共福利中人工智能的负责任使用计划》中概述的类别（例如，信息技术就绪程度和互操作性最佳实践）在公众服务提供生态系统中的

实施方案。

- 探索扩展为州、地方和部落政府提供的采购指南（超出《公共福利中人工智能的负责任使用计划》中提供的内容），以便他们在评估用于管理公共福利项目的信息系统的人工智能工具时使用。
- **HHS 长期优先事项：**
 - 探索为在公众服务生态系统中工作的政府、非营利组织和研究合作机构提供资源，以使用人工智能来改善他们的项目和福利。
 - 确定公众服务组织可以利用的最佳实践开源人工智能和基础设施工具，并将这些工具与特定的高价值用例进行映射。
 - 评估实现 HHS 信息技术基础设施现代化的机会，以支持在公众服务生态系统中更广泛地使用人工智能。
 - 将资源和技术援助机会整合到诸如整笔赠款、高级规划文件、挑战赠款和联邦合同等机制中，用于解决依赖资源的公众服务项目的人工智能应用（例如，促进健康和福祉），并在最合适和可行的情况下实施。
 - 考虑设计类似于美国医疗卫生信息与管理系统学会（CMMI）、美国总务管理局（GSA）和美国国防高级研究计划局（DARPA）使用的“登月”竞赛，以寻求系统级的公众服务交付解决方案，并为有前景的解决方案提供资源和援助以实现规模化。

2. 确保人工智能应用的数据质量和可用性

- **背景：**部分由于遗留信息技术系统、项目要求、对参与者隐私的担忧以及员工 / 客户的数字素养（以及其他因素），许多公众服务机构以非结构化、非标准化的格式记录数据。这些数据难以整合到人工智能驱动的应用程序中。通过提高数据质量、治理水平和互操作性，HHS 可以推动更多需要可访问数据的人工智能用例的应用。此外，人工智能本身可以通过更好地解释非结构化信息来提高数据的可用性。数据质量和可用性的提高可能会增强服务协调的互操作性，并使 HHS 更接近其人本中心的目标，即无缝连接参与者的平台与公众服务、医疗保健和其他公共福利项目。
- **HHS 已采取的行动（非详尽列举）：**《公共福利中人工智能的负责任使用计划》(HHS) 为州、地方和部落政府有效使用人工智能推荐了推动因素，包括提高数据质量和可访问性。
- **HHS 近期优先事项：**

- 继续发布指导方针，并在授权的情况下，建立跨项目、部门、政府层级和社区组织共享数据的互操作性标准。
- 在州、地方和社区组织的参与下，确定公众服务提供中数据质量和收集存在差距的优先领域（例如，针对不太常用语言的翻译），并就改进路径达成一致。
- 推广在公众服务生态系统或相关领域观察到的最佳实践的数据质量标准、治理和访问方式，并根据公众服务的需求进行调整，包括利用人工智能改进数据处理和结构化的最佳实践。
- 探索与私营部门的合作，这些合作可以为有兴趣提高数据质量的 HHS、州、地方和社区组织提供技术援助。
- **HHS 长期优先事项：**
 - 考虑实施共享沙盒环境，以加速用例的试点，并降低了解人工智能应用投资回报率的成本。
 - 审查 HHS 拥有的数据集的质量和対人工智能用例的适用性，并在必要时制定改进计划。

4.6.2 促进可信赖的人工智能开发以及合乎道德和责任的使用

HHS 可以通过以下方式确保人工智能的使用值得信赖且安全：

1. 提供有关平衡对服务人群的风险与人工智能应用机会的指导方针，并建立参与者信任

- **背景：**许多利益相关者积极发声，确保 HHS 服务的人群直接参与人工智能应用的开发以及人工智能模型中数据的确定过程。根据参与者的需求和文化背景定制公众服务中的人工智能，可以提高服务质量和可及性。这对于在数据中历史上代表性不足或被错误代表的人群（例如难民、部落社区和残疾人）以及可能影响人们权利和安全的人工智能应用（例如福利资格确定）尤为重要。关于人工智能的初步研究表明，文化背景会影响人们使用人工智能的偏好。然而，要完全消除对在人工智能应用中错误代表服务群体的担忧，还需要更多的时间和投资。

此外，对参与者数据隐私和安全的担忧可能会抑制技术的使用。例如，人工智能供应商广泛知晓联邦人工智能风险框架并支持相关监管措施，但许多产品并未提供足够的透明度来缓解公众服务利益相关者的担忧。在没有足够指导方针的情况下，试图解决数据偏见和隐私问题的公众服务机构可能会拒绝那些不愿提供其训练数据额外透明度的人工智能解决方案供应商。

最后，HHS 在公众服务提供中执行人工智能使用的现有权力有限。它缺

乏对人工智能的强有力监督方法，包括认证、隐私和安全控制以及第三方评估。为在公众服务环境中开发和部署人工智能的组织制定明确的风险评估和缓解标准，可以降低不当使用的风险。

○ **HHS 已采取的行动（非详尽列举）：**

- 儿童与家庭管理局 2024 年 7 月发布的《生成式人工智能工具政策》提供了鼓励适当和负责任使用生成式人工智能的原则，以支持其员工队伍并改善服务交付。这些要求适用于儿童与家庭管理局的工作人员和承包商，包括了解工具的目的和局限性、了解如何安全使用和保护参与者数据、审查和核实输出内容，并在使用过程中保持透明。
- 《公共福利中人工智能的负责任使用计划》(HHS)：
 - 提供了《管理自动化算法系统使用风险》的建议，包括专注于管理最高风险的人工智能用例、确保安全和保障、维持人类判断力、允许参与者选择退出、保护受助者利益以及维护公民自由。
 - 建议根据美国国家标准与技术研究院 (NIST) 人工智能风险管理框架中概述的六项原则，建立有效的人工智能风险治理机制。其他建议包括维护自动化和基于算法的技术清单、创建正式流程以评估人工智能使用中的风险，以及对供应商进行人工智能治理实践方面的教育。
- 《HHS 值得信赖的人工智能手册 (2021 年)》概述了 HHS 内部使用人工智能应用的指导方针。该手册提供了确保 HHS 内部人工智能应用在整个生命周期中以符合联邦标准的道德、有效和安全的方式开发和部署，并促进公众信任的指导方针。然而，这些指导方针尚未专门针对公众服务项目进行定制，并且大多局限于内部用例的指导，而不是促进外部使用。
- 《关于在自动化系统中执行民权、公平竞争、消费者保护和平等机会法律的联合声明 (2024 年 4 月)》阐明了对自动化系统违规行为采取执法行动的能力，并可利用该权力执行与公平相关的规则，包括执行民权、公平竞争和消费者保护方面的规则。
- 发布内部治理文件供外部参考。儿童与家庭管理局已发布其《人工智能激活工具包》。

○ **其他联邦机构已采取的行动（非详尽列举）：**美国农业部的《州、地方和部落政府使用人工智能的框架》包含了 HHS《公共福利中人工智能的负

责任使用计划》中的大多数风险管理和治理建议。这反映了许多州、地方公众服务部门在管理 HHS 和美国农业部项目（例如贫困家庭临时援助计划和补充营养援助计划）时的重叠职责。

○ **HHS 近期优先事项：**

- 发布《公共福利中人工智能的负责任使用计划》中概述的新指导方针和建议，包括阐明人类在自动化系统中的干预角色、客户支持以及生成式人工智能使用的原则。
- 考虑发布关于在公众服务项目中与参与者进行最佳实践互动的指导方针，以解释并建立对使用人工智能的信任。
- 研究在使用最佳实践安全标准（例如偏差缓解和保持人为干预）的同时，在公众服务中使用人工智能的有效方法。
- 定义在公众服务提供中使用人工智能的适用监管权力（例如，针对辅助生活和社区生活中的人工智能设备），并阐明 HHS 在确保人工智能使用的可信赖性和安全性方面的执法作用。

○ **HHS 长期优先事项：**

- 通过遵守整笔赠款条件、州计划、高级规划文件、挑战赠款和联邦合同的最佳实践指导方针，将人工智能安全和透明度要求纳入 HHS 的资金机制，并与相关联邦合作伙伴进行协调。
- 探索为 HHS、州、地方和社区组织提供直接资源支持的机会，以监测人工智能应用和使用风险。

4.6.3 在美国普及人工智能技术和资源的获取，包括为代表性不足的人群

HHS 可以通过与以下机会相关的行动来确保公平获取人工智能：

1. 提高民众的数字素养和数字普及率

- **背景：** 在一些项目中，可能使用公众服务项目的人群年龄较大且英语熟练程度较低。从历史上看，这些人群的数字素养、互联网接入和智能手机普及率较低。此外，2400 万美国人（其中一些与公众服务人群有重叠）无法接入宽带互联网。这种“数字鸿沟”限制了面向客户的人工智能应用在公众服务中的有效性和解决方案范围。人工智能应用可以缓解这些挑战，但这需要公众服务生态系统的某些部分可能不具备的基本数字素养和能力。

数字鸿沟的另一个后果涉及数据访问和同意。一个机构或社区组织可能无法获取数字接入有限或无法或不愿同意共享其数据的人群的数据。数据差距会加剧数据质量问题，并阻碍公平和情境化的预测分析的部署以

及更好的人工智能工具的开发。提高潜在参与者的连接性和数字体验，可以增加他们对原本因缺乏技术或不了解技术而无法获得的服务的访问。

- **HHS 已采取的行动（非详尽列举）：**《公共福利中人工智能的负责任使用计划》与公众互动，就人工智能在公共福利中的使用收集广泛反馈。这些互动包括听证会、咨询委员会、部落协商、网络研讨会、讲习班和其他活动，旨在让更多人参与 HHS 与人工智能相关的政策制定。
- **其他联邦机构已采取的行动（非详尽列举）：**美国农业部的《州、地方和部落政府使用人工智能的框架》与 HHS 计划中的建议密切呼应，以确保在公共系统中使用人工智能时的公平获取和防止偏见。
- **HHS 近期优先事项：**
 - 建立持续的咨询渠道，包括信息技术 / 人工智能合作组织、社区团体、人工智能主题专家、研究机构、一线工作人员，以及不同人群（例如城市 / 农村、儿童 / 成人、老年人、残疾人）和背景（例如种族、性取向、族裔和语言）的参与者及代表，以确定提高公众服务中人工智能可及性的途径。
 - 制定指导方针，指导州、地方和社区组织如何解决其服务人群中数字素养的不平等问题，包括确定导致人工智能使用差异的不平等现象的历史、当代和结构性因素的指导方针。
 - 分享 HHS 实施的人工智能用例信息，为公众服务组织树立榜样。
 - 汇编并研究潜在的人工智能用例，以缓解或解决不平等问题。
- **HHS 长期优先事项：**
 - 在获得授权并具备资源的情况下，考虑设立赠款或援助计划，以解决公众服务生态系统中获取有影响力的人工智能的不平等问题（例如，针对数字鸿沟较大的人群的奖励）。
 - 与相关实体协调，探索为高风险人群开展教育宣传活动，主题为防范人工智能驱动的诈骗（例如，深度伪造技术支持的诈骗，如冒充家庭成员、政府官员和金融机构），这些诈骗旨在骗取个人的金钱和资源。
 - 继续评估并支持确保服务不足人群能够获取并受益于人工智能的方法。

2. 确定跨部门合作领域以提高人工智能相关的规模经济效益

- **背景：**即使机构和组织有资金投资于自身的人工智能应用，他们可能仍

面临两个额外的获取障碍。首先，较小的组织可能会发现，在缺乏超出其能力范围的更广泛的行业投资的情况下，难以大规模获取人工智能的益处。例如，像欺诈检测或项目测量分析这样的用例可能需要多个组织的数据或技术能力。其次，小型组织可能缺乏内部具备足够技术专长的员工队伍来评估供应商选项并集成新的解决方案（有关与劳动力相关的机会的更多详细信息，请参阅下一节“培养人工智能赋能的劳动力和组织文化”）。因此，这些组织的规模和构成可能会阻碍其获取人工智能，即使存在投资。解决规模挑战可以通过更广泛地获取增强项目的人工智能来改善服务体验和质量。同样，这可以提高服务的可及性，使资源不足的州、地方和社区组织能够通过原本因规模问题无法使用的人工智能平台接触到更多的人。

○ **HHS 近期优先事项：**

- 确定对公众服务提供生态系统最有前景，但需要超出社区组织或州、地方政府规模的用例或信息技术投资（例如，欺诈检测能力）。
- 考虑为部署需要协调活动或更大规模的用例提供赠款和技术援助机会。

○ **HHS 长期优先事项：**

- 探索创建一个“公众服务人工智能”工具包，其中包含公众服务中人工智能应用的关键资源，并将其开源提供给州、地方政府和社区行动组织。
- 考虑召集一个 HHS 人工智能卓越中心团队，为 HHS、州、地方政府和社区组织提供技术专长和能力，并为实现本计划的目标培养他们的能力。

4.6.4 培养人工智能赋能的劳动力和组织文化

HHS 可以通过与以下机会相关的行动来培养人工智能赋能的劳动力和组织文化：

1. 提高公共服务员工的数字素养、人才水平和对使用技术的开放态度

- **背景：**通过与公共服务利益相关者的非正式对话，HHS 了解到他们既对人工智能对其员工队伍的影响表示担忧，又请求在人工智能教育方面提供帮助。这些担忧反映了公共部门普遍缺乏人工智能专业知识，这可能会阻碍人工智能的使用。

此外，在回应 HHS 关于人工智能在公众服务和医疗服务提供中的信息请求时，人工智能开发者和实施者经常将组织准备情况（人员能力、技术基础设施、数据质量、变革管理）列为成功使用人工智能的关键障碍。此外，供应商需要公共服务领域中了解人工智能的专家提供关于以任务为导向的用例

识别和优先级排序的指导方针，但在许多机构和社区组织中这类专家稀缺。此外，利益相关者内部数字人才有限，这阻碍了对新工具的使用和热情。这种差距甚至延伸到从事公众服务提供的非营利组织的最高层，这些组织的董事会通常缺乏具有深厚技术经验的成员。最后，即使机构或社区组织可能对信息技术进行大规模投资，但如果缺乏对人工智能工具的培训，员工可能无法有效地使用新技术，从而可能降低服务质量。当公众服务利益相关者愿意使用新技术时，他们可以利用人工智能来增强员工队伍的能力，从而有可能释放更多资源，为越来越多使用公众服务项目的人群提供服务。

- **HHS 已采取的行动（非详尽列举）：**

- 《公共福利中人工智能的负责任使用计划》为州、地方和部落政府提供了支持员工合理使用人工智能的行动建议，包括培训他们开发和使用自动化及算法系统、在使用人工智能时保持员工的判断力，以及在与第三方供应商合作时对算法系统进行控制。美国农业部的《州、地方和部落政府使用人工智能的框架》与这些建议高度相似。
- 《HHS 值得信赖的人工智能手册（2021 年）》提供了关于 HHS 内部系统中人工智能的教育信息，包括人工智能的益处、缺点以及在部门内的潜在应用。该手册还提供了将值得信赖的人工智能原则纳入工作流程和监督人工智能相关项目的指导方针。

- **HHS 近期优先事项：**

- 探索为 HHS 内部、州、地方和部落政府以及社区组织提供劳动力培训和技术援助的直接赠款或技术援助机会。
- 制定联邦和州机构以及社区组织如何提高其员工队伍人工智能准备程度的最佳实践指导方针。
- 为 HHS 从事公众服务工作的员工开展数字素养和人工智能素养培训。
- 支持或发起公众服务生态系统与人工智能和数字转型领域的私营部门领导者之间的合作，以促进信息共享。
- 在部落国家有需求且资源允许的情况下，支持部落国家对公众服务中人工智能的监管和监督工作。

- **HHS 长期优先事项：**

- 将 HHS 内部的数字和人工智能素养培训资源公开提供给州、地方和部落政府以及社区组织。
- 定期举办公众服务人工智能会议，设置学习课程、实践工作坊并提供录制资源。

2. 利用人工智能缓解公众服务劳动力短缺问题

- **背景：**如前所述，美国劳工统计局预计未来十年美国每年将短缺 67,300 名社会工作者。一个更加数字化、由人工智能赋能的员工队伍能够识别并应用增强公众服务劳动力能力的用例，使员工能够专注于增值活动和与参与者的互动。这可能缓解导致员工满意度低和离职率高的工作压力。然而，使用人工智能来扩充公众服务劳动力也存在一些担忧。首先，人们担心人工智能的使用可能会导致劳动力被取代，但却无法提高生产力或服务质量。其次，过度依赖人工智能来增加劳动力容量可能会使公众服务项目失去人性化的一面。HHS 正在考虑如何在平衡这些担忧的同时，充分发挥人工智能在公众服务劳动力中的作用。
- **其他联邦机构已开展的活动（非详尽列举）：**美国劳工部于 2024 年 10 月发布了《综合人工智能实践指南》，提供了人工智能如何使工人和企业受益的策略，同时关注工人的权利、工作质量、福祉、隐私和经济安全。
- **HHS 近期优先事项：**
 - 分享公众服务提供生态系统中扩大劳动力人工智能能力的最佳实践。
 - 探索发布更多针对公众服务领域的指导方针，以确保人工智能的合理使用符合美国劳工部的人工智能实践指南，并满足参与者对保持人际互动的需求。
- **HHS 长期优先事项：**
 - 审查现有的为州、地方和部落政府提供的项目交付和互操作性指导方针，确定人工智能可以缓解劳动力容量限制的领域。

4.7 结论

人工智能有潜力解决公众服务生态系统中存在的根本性挑战，从长期的劳动力短缺到参与者对项目的满意度较低等问题。最终，人工智能应用可能会改善参与公众服务项目人群的体验。然而，一些根本性挑战阻碍了人工智能的应用，包括资金短缺以及对权利和安全的担忧。尽管存在这些挑战，HHS 相信人工智能可以提高项目质量、增加服务可及性、减轻行政负担，并增强公共福利系统的互操作性。

此外，HHS 处于有利地位，可以帮助公众服务生态系统克服挑战，实现人工智能的益处。同时，HHS 可以制定标准并向公众普及人工智能固有的风险，确保人工智能应用是值得信赖且安全的。HHS 还可以作为召集者，推广来自更广泛生态系统的最佳实践，并突出人工智能的实际应用经验及其对服务人群和历史代表性不足群体的影响。随着人工智能领域的发展，本计划将不断演变，但 HHS 相信本计划中概述的行动将切实推动 HHS

和美国在人工智能领域的战略利益。

5 公共卫生

5.1 引言与背景

在本战略规划中，公共卫生被定义为“通过社会、组织、公共和私人团体以及个人的有组织努力和明智选择，预防疾病、延长寿命和促进健康的科学与艺术”。美国的公共卫生涵盖了广泛的问题，如传染病、药物使用障碍、非传染性疾病、环境健康与气候适应，以及心理和行为健康。公共卫生挑战及其潜在的疾病过程非常复杂，通常涉及生物、社会、经济和其他因素之间的相互作用，需要公共和私营部门的高层参与者之间进行协作。

COVID - 19 大流行及其后续影响凸显了美国公共卫生生态系统中存在的严峻挑战和差距，包括：（1）难以快速收集、共享和分析信息；（2）健康不平等加剧以及公众对科学的不信任；（3）长期的资源和人员紧张问题。

人工智能可以为解决这些挑战提供新的解决方案。例如，人工智能可以通过自动化数据生命周期中的各个流程（如数据清理、验证和聚合），并分析大量数据以识别模式和生成见解，从而改善公共卫生决策、干预措施和项目，确保资源分配到最需要的地方。将人工智能整合到公共卫生领域，有可能显著加强疾病监测和干预措施设计（例如，通过自动疫情检测以及对非传染性疾病的大型数据集和非传统数据源进行快速分析）。此外，人工智能可以帮助提高诊断准确性，更好地吸引不同人群，并优化公共卫生领域通常有限资源的利用。HHS 下属机构，如美国疾病控制与预防中心（CDC）、美国国立卫生研究院（NIH）和美国助理部长准备与响应办公室（ASPR）的战略重点和资源投入，对于推动公共部门的这一变革性变化至关重要。HHS 在推动公共卫生领域的创新方面面临着独特的挑战和机遇，并进而影响私营部门的医疗保健。

近十年来，整个公共卫生生态系统一直在进行与数据现代化和人工智能应用相关的联邦活动。例如，第 13994 号行政命令、第 13960 号行政命令，以及先进战略技术计划（ASTP）的 HTI - 1 和 HTI - 2 规则。这些活动还包括旨在提高数据可用性和质量的举措、各机构的具体实施工作，以及与数据驱动的 COVID - 19 应对措施和未来事件响应准备相关的联邦规则和政策。这些努力虽然并非专门为了快速应用人工智能，但直接关联并支持公共卫生机构部署人工智能的工作。美国公共卫生生态系统的强度取决于其最薄弱的环节；如果没有数据现代化和互操作性，孤立的卫生实体将无法为共享数据和人工智能的使用做出贡献，也无法从中受益。这阻碍了整个生态系统构建有效检测、理解和解决公共卫生问题所需的全面数据视图。目前正在为打破数据孤岛和鼓励人工智能使用奠定基础，CDC 和 HHS 的其他部门有机会成为生态系统中其他参与者的灯塔。

在过去几年中，几乎每个全球卫生论坛都讨论了人工智能创新和应用，人工智能有机会改善全球健康状况。HHS 已经发起了多项与全球健康相关的行动（例如，ARPA - H 的人工智能抗生素项目，以对抗抗菌素耐药性）。然而，正如 HHS 代表在七国集团会议上所指出的，人工智能的有效性在很大程度上取决于一个国家的有利环境——一个值得

信赖、可访问且无偏见的环境。包括美国在内的世界各国都在研究如何最好地利用人工智能来改善医疗保健系统，并防范任何地方出现的重大健康威胁。

除了众多机遇之外，人工智能的使用也带来了与隐私、伦理和公平性相关的严重风险，其中许多风险可以通过 HHS 的行动进一步解决。为了最大限度地发挥人工智能在公共卫生中的益处，必须加快现有努力，并将其整合到一个平衡创新与安全保障的统一战略中。为此，在本文件的后续部分，HHS 概述了一系列战略重点，以促进健康人工智能的创新与应用，确保人工智能的使用值得信赖且安全，普及人工智能技术和知识的获取，并支持培养人工智能赋能的劳动力和组织文化。

5.1.1 行动计划摘要

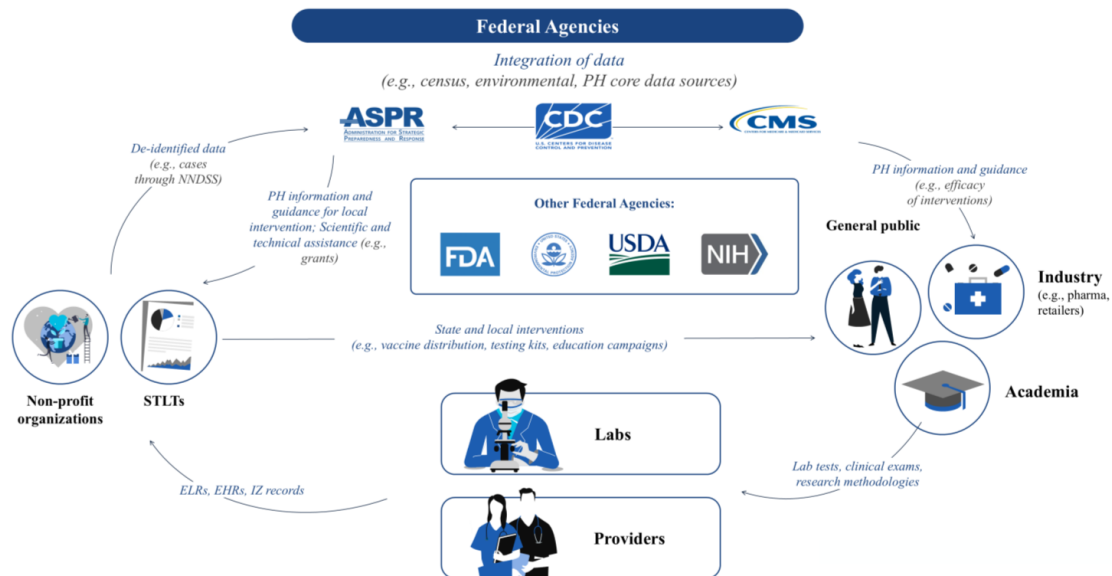
在本章的后续部分，HHS 将阐述为推进其在公共卫生领域负责任使用人工智能的四个目标而提议采取的行动。以下是每个目标下行动主题的摘要。有关提议行动的详细信息，请参阅 5.6 行动计划部分。

关键目标	提议行动主题（非详尽，详见 5.6 行动计划）
1. 催化健康人工智能创新与应用	鼓励研究，制定指导方针，确定支持证据生成和人工智能在公共卫生领域规模化应用的资源；实现实施人工智能和支持应用所需的基础设施现代化
2. 促进可信赖的人工智能开发以及合乎道德和负责任的使用	建立保障措施，确保数据质量和准确性；在公共卫生生态系统中标准化数据安全政策；开发并推进从始至终考虑和评估健康公平性的人工智能工具和技术
3. 普及人工智能技术和资源	营造一个能够在公共卫生生态系统中实现数据共享的环境；支持人工智能的应用、开发和合作，特别是为资源有限的州、地方和部落政府以及社区组织提供支持；开发用户友好、可定制的开源人工智能工具，以扩大获取途径并满足不同用户的需求
4. 培养人工智能赋能的劳动力和组织文化	扩充和支持公共卫生劳动力，以应对职业倦怠和人员流失问题；推广适合每个社区独特需求的人工智能教育和基于社区的人工智能方法

5.2 参与公共卫生人工智能价值链的利益相关者

美国公共卫生生态系统依赖联邦政府和州、地方、部落政府（STLTs）的协调与支持，并依靠从医疗服务提供者、医疗系统、私营合作伙伴、研究人员到非营利组织和公众等广泛利益相关者的协作，以实现积极的社会变革。为了说明公共卫生参与者的多样性，以下是一个非详尽的、具有代表性的利益相关者之间示例流程图表（图 12）和涉及的利益相关者列表。请注意，该图表和列表并未涵盖所有利益相关者的角色和互动。有关监管指导和权限的更多详细信息，请参考其他 HHS 文件。

NON-EXHAUSTIVE | ILLUSTRATIVE



- **HHS 机构：**公共卫生通过 HHS 各运营部门的努力得到支持，例如：
 - **ACF：**为家庭和儿童提供福利和服务，其中许多与公共卫生相关（例如，行为健康和预防虐待）。
 - **ACL：**支持针对有复杂需求人群的项目，特别是老年人和残疾人，包括营养服务、医疗护理和老年人支持服务。
 - **AHRQ：**致力于通过研究提高所有美国人医疗保健的质量、安全性、效率和有效性。
 - **ASPR：**领导国家对灾害和公共卫生紧急情况的准备、响应和恢复工作。
 - **有毒物质和疾病登记署（ATSDR）：**防止人们接触有害物质（如化学品、农药、重金属），并减轻相关的健康风险。ATSDR 进行风险评估和健康咨询，并支持健康教育。
 - **CDC：**积极检测、监测、定义、预防和应对疾病爆发，管理国家卫生项目，并通过提供技术援助和信息支持政策制定。
 - **CMS：**管理主要的公共医疗保健支付项目（如医疗保险和医疗补助），制定这些项目中医疗服务提供者的参与条件，要求他们共享关键的公共卫

生数据（例如，与医疗相关感染的数据），并可为特定的设备或服务提供支付。

- **FDA**：作为核心监管机构，确保医疗产品的安全性和有效性以及医疗设备（包括人工智能医疗设备）的安全性。
 - **IHS**：为美国印第安人和阿拉斯加原住民提供全面的医疗服务体系，确保提供符合文化背景的公共卫生和公众服务，以提高该人群的身体、心理、社会和精神健康水平。
 - **HRSA**：旨在改善未参保、孤立或高风险人群的医疗保健获取途径。
 - **NIH**：作为美国生物医学和行为研究的管理者，通过维护健康数据存储库（如美国国立医学图书馆数字序列信息）和开展公众宣传活动，支持公共卫生工作，以促进明智的健康决策。
 - **SAMHSA**：领导公共卫生工作，以促进国家的行为健康，改善患有精神和物质使用障碍的个人及其家庭的生活。
- **公众**：普通民众通过参与预防措施和其他行动，在公共卫生中发挥着至关重要的作用，这些行动包括接种疫苗、保持卫生、个人健康和生活方式选择，以及疾病和症状报告。这包括接受服务的个人及其护理人员。
 - **其他联邦机构**：HHS 以外的联邦机构，如环境保护署（EPA）和教育部（ED），是支持公共卫生行动的关键合作伙伴和数据提供者。这包括通过公共福利项目提供的信息、人口数据、环境数据以及就业和经济数据。
 - **公共卫生服务军官团**：美国公共卫生服务军官团体现了联邦机构之间的跨机构合作和价值。军官团的成员服务于 21 个联邦机构，展示了所有机构在负责任地使用人工智能以支持公共卫生方面所发挥的重要战略作用。
 - **STLTs**：STLTs 和自由联合州的卫生部门是公共卫生生态系统的支柱，是 HHS 在公共卫生工作中的关键合作伙伴。STLTs 负责其社区的健康和福祉，管理与联邦卫生机构共享的数据集，支持社区内的预防和干预措施（例如疫苗分发），并向各种当地利益相关者和组织发布指导方针。
 - **公共教育和宣传组织**：沟通和公共教育活动是公共卫生价值链的关键组成部分，包括推广免疫接种活动。有几个实体（例如医疗保险 PACE 项目）代表美国卫生局局长和公共卫生服务军官团，其使命是进行公共卫生教育和宣传，例如推广免疫接种活动。这些组织可能会获得与特定活动直接相关的联邦资金、私营合作伙伴的资金，或当地医疗保健组织的资金。医疗服务提供者、支付者和社区组织也发挥着关键作用。
 - **学术界和研究机构**：学术和研究组织，包括相关的医院、实验室和研究机构，是

科学研究的主要生产者。它们为下一代公共卫生工作人员提供培训，并作为该领域创新的关键枢纽，特别是在人工智能用例方面。

- **医疗系统、医疗服务提供者和实验室：**医疗系统对于成功实施持续的和紧急的公共卫生项目至关重要，并通过健康登记、监测系统和研究数据库生成数据，为政策制定提供信息。
- **制药、生物技术和医疗设备行业：**私营生命科学组织通过大规模提供和分发药物、生物制品和医疗设备，支持公共卫生工作。它们还包括参与医学研究和发现的研究人员和主题专家，是人工智能创新的主要来源。
- **全球合作伙伴：**全球合作伙伴，包括多边组织、双边组织、非政府组织、外国政府和其他机构，与美国公共卫生机构合作，应对跨越国界的健康挑战。它们的共同行动有助于促进知识和数据的共享，支持早期缓解传染病，预防公共卫生紧急情况，支持医疗保健系统的能力建设，并确保各地区公平获得医疗服务和干预措施。
- **非营利组织和社区组织：**全国性的公共卫生合作组织，其成员通常由致力于特定公共卫生职能（如流行病学家）或目的（如加强公共卫生实验室）的人员和实体组成，作为联邦和 STLT 公共卫生机构的合作伙伴、渠道和实施中介发挥着重要作用。此外，融入社区的非政府组织支持提供健康和福祉服务，以确保公共卫生项目有效地覆盖弱势群体。
- **基金会和私人资助者：**基金会可以通过为社会决定健康因素（SDOH）等领域的临床试验和研究提供资金，或直接提供公共卫生服务，来支持公共卫生工作。此外，其他资助者可能会投资于价值链中的组织，如技术公司。
- **技术公司：**这些公司包括专注于人工智能基础设施（如云存储）的公司、大型多元化科技公司、数字解决方案供应商和白帽黑客。这些公司为利益相关者使用人工智能提供基础设施和服务。

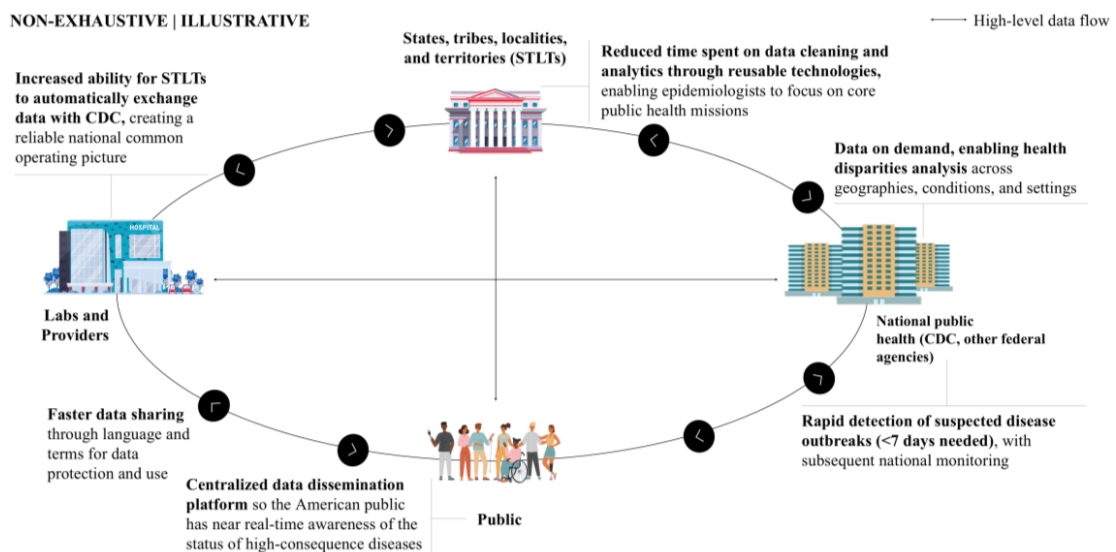
HHS 的几个部门（如先进战略技术计划、政府间和对外事务办公室、美国国立卫生研究院等）通过双边和多边合作、会议以及跨国组织，推进全球健康人工智能工作。例如，全球数字健康伙伴关系是世界卫生组织与各国政府之间的合作，旨在支持全球数字健康服务的实施。此外，HHS 各部门的主题专家作为代表，向七国集团（G7）、二十国集团（G20）和经济合作与发展组织（OECD）等跨国组织提供政策投入和反馈。

数据价值链

上述利益相关者在公共卫生中扮演着许多关键角色，包括数据收集。如前所述，如果没有高质量的数据、适当的数据收集和标准化，人工智能产生见解的能力可能会受到限制。以下部分概述了公共卫生中的数据价值链、CDC 现有的数据改进工作，以及加强公共卫生数据生态系统的其他行动。

作为公共卫生数据流程中的核心参与者（图 13），CDC 已经通过其数据现代化倡议（DMI）取得了重大进展。该倡议致力于提高数据的可访问性、及时性和全面性，以支持日常公共卫生应对工作。DMI 旨在解决公共卫生职能方面的问题，如提高数据共享速度（例如，通过制定数据保护和使用的语言和术语）、增强 STLts 与 CDC 交换数据的能力（例如，通过自动数据管道），以及实现疾病的近实时公开报告（例如，通过集中式数据传播平台）。这些旨在在公共卫生各级建立统一数据管理方法的投入，可以为支持更多人工智能用例奠定基础。

图 13: CDC DMI 数据流



DMI 近期的重要成就包括：

- 通过将当前数据基础设施与通过可信交换框架和常见协议（TEFCA）交换信息的要求相结合，连接公共卫生和医疗保健系统，支持使用互操作性标准，如美国通用临床数据交互规范（USCDI）和美国通用临床数据交互规范 + 倡议，并使用中介减少点对点连接（图 13）。例如，2023 年，CDC 帮助 90% 的流行病学和实验室能力项目受助方连接到公共卫生实验室协会信息服务、ReportStream 或健康信息交换平台，以获取实验室数据。
- 通过支持实施自动双向电子报告馈送（如电子病例报告（eCR）、电子实验室报告和入院 - 出院 - 转院馈送），实现数据访问的自动化和改进，以减少手动报告。2023 年，CDC 帮助 34 个辖区实施 eCR 数据，以改善病例监测。
- 简化数据收集和处理流程，确保数据一次收集并在公共卫生实体间重复使用，减少重复工作并提高集成度。CDC 正在向集成云计算数据平台迁移，并在关键数据系统中使用人工智能用例（例如，现代化国家生命统计系统，以自动对多种死因进行编码）。
- 实施核心数据使用协议（DUA），以统一和增强全国各辖区之间的数据交换。

继续这些努力，并在公共卫生、医疗保健（特别是初级保健）和社区组织之间进行更多整合，可以进一步增强韧性，并在应急响应和日常环境中改善医疗服务（例如，通过医疗保健系统电子健康记录与社区组织的本地人口统计数据之间的自动数据交换）。有一个大胆的机会可以在 CDC 和其他机构努力的基础上，进一步整合数据，这既可以得到人工智能的支持，又能使人工智能的应用推进公共卫生优先事项。

5.3 人工智能在公共卫生中的应用机会

随着人工智能技术的日益普及，HHS 将努力确保公共卫生合作伙伴以合乎道德、可靠和公平的方式将人工智能整合到公共卫生组织和使命中。人工智能可以在多个领域为公共卫生优先事项和基础设施提供支持：

1. **改进威胁检测、数据驱动决策和干预措施的有效性：**有机会利用人工智能聚合和分析更大、更复杂或非结构化的健康数据集，包括医疗服务提供数据（如索赔或电子健康记录数据）以及非健康数据集（如移民模式和气候数据），这可以支持更及时有效的干预措施。例如，将二次数据整合到监测系统中，以更好地预测和应对新出现的公共卫生威胁；整合社会决定健康因素（SDOH）和其他数据集，以更好地理解非传染性疾病（如糖尿病和心血管疾病）的潜在风险因素和疾病过程，从而为有效的干预措施设计提供信息；更广泛的应用（如数据探索）可以通过支持跨疾病领域的初始研究综合，在适当的人为监督和透明度下，加速指南的制定。然而，在公共卫生中部署人工智能需要制定和使用指南，以解决相关的伦理和安全问题。
2. **优化有限资源的分配，特别是在公共卫生紧急情况下：**资源问题长期以来一直是美国公共卫生系统面临的挑战，资金可能分散且有时不稳定。为了更好地确定优先事项，以实现最大的健康结果和公平性，人工智能可用于确定现有干预措施能产生最大影响的高风险地区，确保正确的资源在正确的时间到达正确的社区。在 COVID - 19 大流行期间，包括美国在内的许多国家优先为一线工作者和老年人等高风险人群分配疫苗，这一工作在某些情况下得到了预测分析的支持。人工智能还可以通过多种方式进一步助力应急响应，例如供应链的预测建模和疫苗可接受性的地图绘制。最近，CDC 一直在使用人工智能为干预措施提供信息，并加速对疫情爆发的响应。人工智能可以在公共卫生紧急情况 and 解决其他领域（如非传染性疾病）的持续项目中支持资源优化。
3. **提高公共卫生运营效率，支持公共卫生工作者更好地服务社区：**在公共卫生生态系统中以负责任、安全和战略性的方式使用人工智能，可以大大减轻面临职业倦怠和过度工作量挑战的医疗保健系统和公共卫生部门的运营负担（例如，2022 年，46% 的卫生工作者经常或非常频繁地感到倦怠）。例如，人工智能可用于自动化与撰写和审查资助申请相关的流程，以降低成本，或用于自动化数据录入和合规报告等耗时活动，但前提是要有足够的人为监督。公共福利项目中的行政负

担估计占总医疗保健支出的 15% 至 30%，其中一半包括可以自动化的常规或重复性任务。目前的人工智能工具并不总是适合这些特定任务，需要进一步开发，以确保在有效性与安全性和准确性之间取得平衡。此外，为了使用这些工具，公共卫生专业人员需要有提升技能和接受培训的机会，以安全有效地使用人工智能。

4. **增强健康公平性，改善服务不足人群的医疗服务可及性：**人工智能应用，特别是生成式人工智能，为改变公共卫生决策和项目的实施方式提供了独特的潜力，尤其是对于传统上服务不足的人群，但前提是要充分防止潜在的偏见。人工智能可以通过消除决策中的人为偏见、进行更有针对性的外展（例如，识别并接触高风险、高需求人群）以及基于证据的个性化信息传递（例如，根据语言需求、健康素养和当地社区背景），来促进健康公平性并改善医疗服务的可及性，从而提高公众对公共卫生指南和项目的认识和接受度。例如，CDC 在 2020 年推出了冠状病毒自我检测聊天机器人，以帮助个人决定是寻求医疗护理还是在家中管理症状。此外，通过自动翻译、转录和个性化等功能，人工智能可以快速生成符合不同人群健康素养、语言和当地背景的内容。人工智能还可用于支持公共卫生工作者或服务接受者的培训和指南制定。

5.4 人工智能在公共卫生中的应用趋势

技术和科学的进步在历史上一直在加速公共卫生的改善，这一现象在最近的 COVID - 19 大流行中在全球范围内凸显出来。全球公共卫生生态系统研发出了安全有效的 COVID - 19 疫苗，并在 2024 年 8 月前实现了全球 70% 以上的接种覆盖率。尽管取得了这一成功，但大流行也暴露了美国公共卫生系统存在的多重挑战（如过时的系统和有限的资源）以及创新的机会。人工智能技术有望缓解未来的公共卫生危机并加强公共卫生系统。值得注意的新兴趋势包括但不限于：

1. **人工智能在公共卫生中的应用既带来热情，也引发担忧：**人们对在医疗保健和公共卫生中使用人工智能的兴趣日益浓厚，但美国公众和卫生官员也对其潜在影响表示担忧。一些公共卫生利益相关者正在迅速使用人工智能，并频繁提及它，只要公平性和伦理问题得到解决，他们就乐于继续推进。例如，早在 2022 年 2 月，就有超过 4500 篇科学论文提及在应对大流行中使用人工智能和机器学习，其中 239 篇涉及监测，219 篇涉及预测。相比之下，在最近的一项全国性调查中，超过一半的成年人不确定人工智能对在线寻求健康信息的人的影响，另有 23% 的人认为人工智能弊大于利。未来需要同时应对人工智能带来的热情和担忧。
2. **预测分析助力早期疾病检测，加速公共卫生响应：**人工智能的实时预测和预测能力正在通过提供实时监测和预测模型，彻底改变流行病学，从而为积极主动的公共卫生响应提供信息。人工智能扩展了可用于预测和跟踪公共卫生威胁的数据类

型库。潜在的例子包括图像识别,以帮助从医学扫描中早期检测疾病;音频分析,通过语音模式监测心理健康;以及从文本健康记录和社交媒体中获取自然语言处理见解,以跟踪疾病传播和公众情绪。在这些领域,人工智能的应用已经取得了强劲的发展势头。

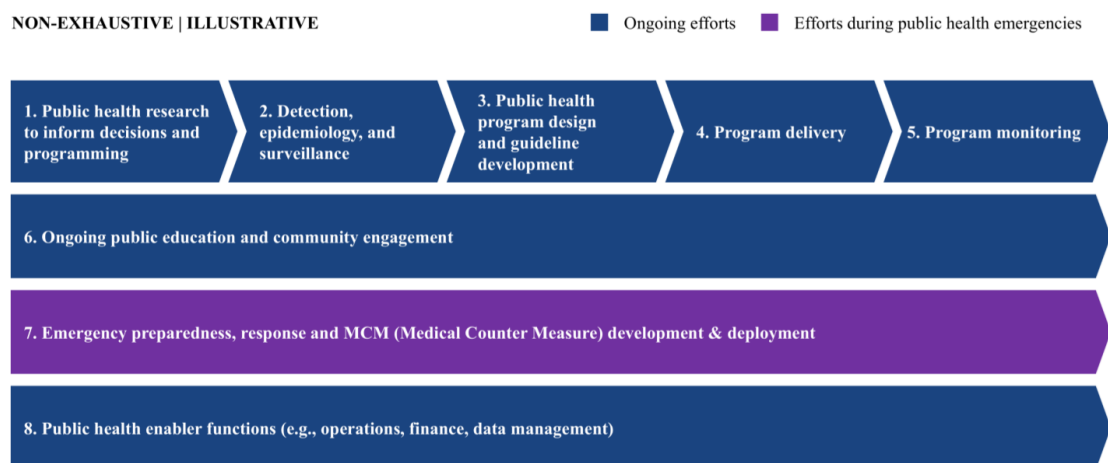
3. **资源限制、基础设施不足、技术知识缺乏和数据匮乏限制了人工智能在公共卫生中的实施:** 传统上服务不足的人群和更脆弱环境中的有限资源可能导致数据匮乏,以及难以建立人工智能所依赖的必要技术基础设施。过时的数据基础设施也可能限制人工智能的使用,尽管 CDC 的 DMI 等举措和对 STLTs 的资助正在帮助改善核心基础设施。此外,公共卫生实体在吸引、留住和培训高质量技术人才方面常常面临挑战,COVID - 19 大流行进一步加剧了这一挑战。
4. **人工智能在公共卫生中的应用和使用情况不一致:** 包括公共卫生在内的公共部门领域在利用人工智能方面存在不均衡的情况,对人工智能的认识和专业知识水平也各不相同。这在很大程度上是由高质量数据的可用性、不同的领域需求、计划中和正在进行的合作努力,以及现代化数据平台和资金的可用性所驱动的。如上文所述,人工智能的广泛潜力值得在更广泛的实施方面进行更大的投资。

5.5 人工智能在公共卫生中的潜在用例和风险

以下价值链虽然并非详尽无遗,但突出了公共卫生的核心运营和项目领域,特别强调了在急性公共卫生紧急情况下的准备和响应。有关相关主题的更多详细信息,特别是产品开发生命周期,请参考医学研究与发现以及医疗产品开发、安全性和有效性章节;有关医疗服务提供的内容,请参考医疗服务提供章节;有关公众服务项目(通常涉及社会决定健康因素)的交付,请参考公众服务章节。

这个框架是对公共卫生人工智能应用多样性的说明性表示,应根据组织运营的具体背景进行调整,包括传染病控制、慢性病预防和管理等领域。

图 14: 公共卫生价值链



公共卫生价值链的每一步都为人工智能提供了提高公共卫生工作效率、增强分析能力和复杂性、改善医疗保健和信息获取，以及提高对公共卫生优先事项认识的机会。

同时，人工智能也伴随着风险，其中许多风险在多个用例中都存在。在公共卫生应用中，一些常见的风险包括偏见（由于训练数据中的缺陷或代表性不足，对某些群体进行有意或无意的歧视）、虚构信息（编造来源或信息）、难以解释（由于数据集庞大和参数众多，难以解释人工智能结果）、拟社会关系（用户因与人工智能模型进行“逼真”交互而产生的社交化解读）以及未经授权披露机密信息。此外，如果没有全面的保障措施，可能会为了节省成本而使用人工智能，即使其效果不如传统技术。

对公共卫生而言，一个特别具有挑战性的人工智能风险是错误信息和虚假信息；正如 COVID - 19 大流行所表明，“根据当时可获得的最佳证据判断为虚假、不准确或具有误导性的信息”现在能够以前所未有的速度和规模传播（例如，通过社交媒体和搜索引擎），并可能导致严重的公共卫生后果，如对卫生工作者的骚扰和暴力、对隔离指南的不充分遵守，以及推广未经证实的医疗治疗方法。此外，鉴于生成式人工智能建立在具有大量参数的神经网络之上，很难解释见解和建议是如何产生的，这有时被称为“黑箱问题”。再加上可能出现的错误回复和逼真的深度伪造（看起来真实的假图像、音频或视频），人工智能可能会影响公众信任，并最终降低公共卫生项目的有效性。

以下将进一步考虑其中一些风险；HHS 将继续支持针对这些风险的缓解措施，与本文档后面讨论的行动计划保持一致。

在下面的表格中，HHS 列出了公共卫生价值链中人工智能的潜在益处和风险，但并非详尽无遗。请注意，以下详细介绍的用例突出了该领域各种利益相关者现有或潜在的人工智能使用方式。有关 HHS 及其部门如何使用人工智能的详细信息，请参考《HHS 人工智能用例清单 2024》。

功能组件 1：公共卫生研究以指导决策和规划

开展研究与分析工作，以了解和提升人群健康状况，并为未来的项目与决策提供依据。

- **潜在用例（非详尽）：**通过对大型、复杂且通常是非结构化数据集的快速分析得出新颖见解，为规划提供信息；人工智能能够实现对以前未使用或未充分利用的数据（超越传统的表格和数字格式，如音频文件和图像）进行更大范围的数据捕获和分析，从而为更有效的干预措施提供信息；人工智能算法可以从非结构化的电子健康记录中提取临床见解，并进行预测和分类任务，这使用传统方法很难做到，进而为公共卫生干预措施和人口研究提供信息；整合多种数据类型和二次数据到现有的公共卫生模型中，为研究和有效的干预措施提供信息，例如整合健康数据集（如病例数据和废水数据）和非健康数据（如移民模式、旅行和销售模式以

及搜索引擎数据), 可以为正在进行的公共卫生优先事项和新出现的威胁的预测和监测提供信息; 使用合成数据和数据链接技术推进研究并保护隐私, 合成数据是人工生成的, 可模拟患者或人口数据而不包含任何实际个人信息, 使研究人员能够进行研究和测试算法, 而不会有暴露个人身份信息 (PII) 或受保护健康信息 (PHI) 的风险, 隐私保护记录链接 (PPRL) 技术进一步增强了这种能力, 它允许在不同数据库中匹配对应同一实体的记录, 例如数字双胞胎和情景建模, 虚拟复制物理系统或过程 (如个性化患者模型) 可用于模拟不同情景 (如公共卫生情景建模和临床试验), 以优化公共卫生干预措施或治疗计划并改善结果。

- **潜在风险 (非详尽):** 可能引入偏见和歧视, 例如由于人工智能通常基于历史数据进行训练, 而这些数据往往更关注特定的人口群体, 导致排除代表性不足的群体, 从而产生不具有代表性的研究结果, 这些结果不能平等地适用于所有群体, 并延续现有的偏见; 可能降低有效性或可解释性, 例如由于多因素数据导致结论不明确, 包含多个变量的大型数据集可能会错误地发现不存在真实联系的关联; 可能过度使用合成数据, 例如随着合成数据的迭代使用, 模型完整性和多样性表示可能会下降, 即使使用合成数据的初衷是增加多样性, 但在分析和再分析以生成更多合成数据的过程中, 可能会损害模型质量, 危及结果的准确性和有效性, 最终无法实现代表不同人群和 / 或减少偏见的潜在目标。

功能组件 2: 检测、流行病学和监测

用于分析疾病趋势、识别疾病爆发, 以及研究人群中健康事件的分布情况和决定因素的数据及模型。

欲了解更多信息, 请参阅“医疗服务提供”以及“医疗产品开发、安全性与有效性”章节中的“潜在用途”部分。

- **潜在用例 (非详尽):** 利用人工智能驱动的传染病监测和预测, 人工智能模型可用于处理大量健康和非健康二次数据集, 以识别潜在的热点地区或疫情爆发, 并监测疾病的持续传播和变化; 开发智能疾病诊断工具以改善临床决策, 人工智能和机器学习算法可用于改善临床决策和从成像系统进行诊断, 以推进对非传染性疾病和正在进行的公共卫生优先事项的检测, 例如人工智能驱动的心力衰竭检测; 推进精准公共卫生以优化资源, 将精准医学 (如基因组学和代谢组学) 与基于人群的策略相结合, 有助于“在正确的时间为正确的人群提供正确的干预措施”, 例如识别高风险地区或人群, 精准公共卫生可以识别弱势社区, 使公共卫生实体能够采取积极主动的行动。
- **潜在风险 (非详尽):** 可能降低可解释性, 例如人工智能模型对小的干扰过于敏感, 可能会错误地将变化报告为公共卫生事件, 从而误导公共卫

生工作；可能泄露机密信息，例如检测算法如果能够访问受保护健康信息（PHI），可能会在输出中无意透露 PHI 或其他识别信息，或者受到网络安全威胁。

功能组件 3：公共卫生项目设计和指南制定

制定相关策略与干预措施，以改善人群健康状况，预防疾病及持续性健康问题（如癌症和糖尿病）。这还包括制定公共卫生指南（如疫苗接种建议以及医疗产品的上市后监测）。

- **潜在用例（非详尽）：**设计超本地化的公共卫生项目以优化资源，人工智能可用于聚合和分析本地健康数据，以更好地了解特定需求（如按社区划分的疾病流行情况）、风险（如环境和社会经济因素）和 / 或基础设施容量（如医疗工作者的可用性、个人防护设备的获取、检测的可及性和营养的获取），从而创建有针对性的项目，优化资源利用，例如众包空气质量分析和倡导，利用人工智能整合社区意见和各种数据源（如市民报告和照片以及排放数据），不仅能够推进科学研究，还能推动社会变革（例如，根据当地污染模式及其对健康的影响的数据，确定具有直接效果的实际行动）；自动化资助申请和提案请求（RFP）的撰写或审查过程以提高效率，在人为监督和透明度的支持下，人工智能可以自动化资助申请和 RFP 撰写或审查过程中的部分手动步骤（例如，汇总数据、校对以确保准确性并符合提交要求），为政府实体和非营利组织带来显著的效率提升，人工智能工具可以快速筛选 RFP 回复或提案，以综合应用中的关键趋势或差距，支持人工主导的审查过程。
- **潜在风险（非详尽）：**可能泄露机密信息，例如算法如果能够访问受保护健康信息（PHI），可能会在输出中无意透露 PHI 或其他识别信息，或者受到网络安全威胁；可能设计出切实际的干预措施，例如人工智能模型专注于人口子集分析，可能会使项目设计过于狭窄，导致创建的高度特定的干预措施无法在社区中推广，这在公共卫生领域是不切实际的结果，因为公共卫生问题普遍存在；可能误解或错误描述资助申请，例如人工智能出现的错误如果违反了管理资助批准或延续的联邦法规，可能会导致资助申请人的正当程序问题，并引发对资金合理分配的担忧。

功能组件 4：项目交付

公共卫生项目的实施与管理

- **潜在用例（非详尽）：**个性化项目交付以增强可及性和公平性，面向公众的人工智能工具可用于在广泛的人群和疾病中高效地提供个性化的健康建议或项目，例如人工智能聊天机器人，聊天机器人应用程序和界面可以根据用户的文本或语音输入进行对话，并生成广泛的非脚本化对话回复，例如 CDC 的 COVID - 19 聊天机器人和世界卫生组织的 S.A.R.A.H

生成式人工智能工具，可根据用户的视频或文本输入提供关于健康主题（如营养和压力管理）的定制信息；使用人工智能预测需要额外资源的地区或人群，考虑到不断变化的因素（如医疗保健使用情况、疾病流行率和资源可用性数据的快速处理），动态调整资源优先级和分配或活动。

- **潜在风险（非详尽）**：可能会使用户困惑或提供不准确的建议，例如用户可能会将人工智能聊天机器人误认为是人类互动，从而产生情感依恋或其他拟社会关系，对心理健康产生不良影响；可能会使劳动力失去动力，例如工作人员可能认为人工智能会取代他们的工作，这可能会降低他们对工作的投入度，进而影响公共卫生工作的质量。

功能组件 5：项目监测

跟踪和评估公共卫生项目的进展、合规性以及成效

- **潜在用例（非详尽）**：改进对数据问题和异常的检测，以提高项目的有效性和效率，基于人工智能的项目监测可以持续、主动地识别数据中的异常值或离群值，这些异常情况可能表明潜在的健康问题或错误。例如，检测数据异常，人工智能模型可用于检测意外的项目结果，这有助于检测不良事件（如设备故障）或异常情况（如不寻常的结果、潜在的欺诈行为），值得进一步调查或调整。
- **潜在风险（非详尽）**：可能存在分析不完整或不正确的情况，例如人工智能驱动的项目分析可能无法充分考虑定性的参与者反馈或影响项目绩效的背景因素（如文化、经济或社会条件），从而导致决策效果不佳。

功能组件 6：持续的公共教育和社区参与

开展双向且持续的工作，旨在向社区及个人传播信息、促进参与并实现合作，以此提升健康教育成效，确保项目的可持续性

- **潜在用例（非详尽）**：个性化公共卫生信息传递和教育，以增加可及性并提高公平性，人工智能可用于为不同的人口群体定制信息，并以低成本将宣传范围扩大到更广泛的受众（例如，全天候可用、自动内容审核）。例如，内容生成工具，人工智能工具可以协助组织完成诸如起草电子邮件活动、创建有吸引力的网页以及个性化吸引特定人群的内容等任务；促进包容性沟通，以提高可及性和公平性，基于人工智能的翻译技术可以帮助弥合语言和 / 或文化差距，使组织能够大规模地接触到多样化的受众。例如，实时翻译应用程序，先进的翻译工具可以实现适应特定语言、方言或行话的实时交互，这有助于建立信任、促进公平并提高参与度。
- **潜在风险（非详尽）**：可能产生不准确的信息传递，例如人工智能用于内

容创建时可能会误解受众，从而传递不适当的内容（如文化上不敏感的建议）或不适当的语气（如对普通公众使用医学术语）；可能传播错误信息和 / 或虚假信息，特别是在不确定或不断变化的情况下，人工智能生成的深度伪造内容和其他虚假内容可能会通过互联网迅速传播，引发公众的不确定性和对现行公共卫生指南的不信任。

功能组件 7：应急准备、响应以及医疗对策开发与部署

预防、发现和应对急性健康威胁的策略与干预措施的设计、协调与实施

- **潜在用例（非详尽）：**支持公共卫生应急人员，以提高他们响应的效率和效果，在公共卫生紧急情况和响应期间，人工智能可用于减轻工作人员面临的即时负担，提高培训和入职计划的效率（例如，根据当地情况定制的医疗工作者培训计划），并支持快速响应（例如，资源分配）。例如，自我学习救援机器人，基于人工智能的机器人系统可以在灾害预防和响应中提供支持（如侦察未知情况、识别危险以及在危险环境中进行救援行动）；传播实时公共卫生指南，以提高可及性，人工智能系统可用于快速整合数据源、生成警报并针对性地分发紧急信息。例如，天气预警信息，人工智能模型可以分析地理、天气和用户数据，生成相关且信息丰富的警报（例如，根据位置以及用户是否正在驾驶，发出不同的冬季天气预警）；开发医疗对策，人工智能有潜力为确定投资方向（例如，分子或疫苗研发管道）提供更明智的决策，并更好地实时监测医疗对策的有效性。例如，识别潜在的药物靶点，这些靶点发现方法可以帮助揭示疾病的新靶点和途径，从而加快干预措施的开发。请注意，医疗对策的开发尤其具有跨领域性，主要在本战略规划の医学研究与发现以及医疗产品开发、安全性和有效性章节中进行讨论；评估和借鉴过去的应急响应工作，以提高干预措施的效果，人工智能工具可用于分析过去公共卫生响应的数据，以识别趋势、成功经验和改进机会（例如，分析各种州级 COVID - 19 政策的影响，为未来的公共卫生决策提供信息）。例如，基于人工智能的事后审查，一个基于人工智能的审查过程可以分析一个组织对紧急情况或灾难的响应，将其与大量以前的记录进行比较，以确定需要改进的领域，并制定有针对性的建议和计划（例如，模拟演练）。
- **潜在风险（非详尽）：**可能误导工作人员，例如在紧急情况下使用的人工智能工具可能会误解紧急情况，为应急支持人员提供不适当的指示，如果这些指示不准确或被误解，可能会造成危害；可能误导资源分配，例如人工智能模型得出的不准确结论（如无效的药物靶点）可能会错误地使人对干预措施产生信心，导致资源分配不当。

功能组件 8：公共卫生支持职能（如运营、财务、信息技术和数据）

支持公共卫生服务提供与管理所需的基础设施和行政流程

- **潜在益处和示例用例（非详尽）：**自动化行政和运营任务，以提高效率，与许多其他组织一样，人工智能为公共卫生机构在结合人为监督的情况下实现运营效率和减少人为错误提供了重要机会，特别是在劳动密集型的行政任务领域。例如，数据录入，由人工智能辅助的更智能的数据录入可以帮助将各种格式的信息转录到集中式数据库中，并通过预测文本字段和实时错误检查算法提高数据的整体质量。
- **潜在风险（非详尽）：**可能使劳动力失去动力，例如那些工作职责涉及可自动化任务的个人可能会认为人工智能的作用是取代他们的职位和职责。

5.6 行动计划

鉴于公共卫生领域人工智能格局的不断演变，美国卫生与公众服务部（HHS）已采取多项措施，推动全生态系统的基础设施更新，并制定相关指导方针，促进人工智能技术得到安全、规范的应用。以下行动计划围绕支持 HHS 人工智能战略的四个目标展开：1. 推动健康领域人工智能创新与应用；2. 促进可信赖的人工智能开发，以及技术的伦理性 and 规范性使用；3. 实现人工智能技术与资源的普及；4. 培育具备人工智能能力的劳动力和组织文化。针对每个目标，行动计划提供了相关背景、HHS 及其他联邦机构截至目前的行动概述，以及 HHS 将采取的具体短期和长期优先事项。HHS 认识到，随着技术的发展，该行动计划需要适时修订，并致力于在提供架构的同时保持灵活性，以确保产生长期影响。

5.6.1 推动健康领域人工智能创新与应用

人工智能的应用与实施有望彻底改变公共卫生领域，并通过例如强化疾病预测等方式，保护公众免受新出现和持续存在的威胁。私营部门在医疗服务或研发领域会深度参与人工智能创新与投资，与之相反的是，它们在公共卫生领域的参与则较为有限。因此，HHS 在公共卫生生态系统中分配资源、协调激励措施以及指导人工智能的实施与应用方面，将发挥更为关键的作用。为此，HHS 可以通过以下方式应对当前创新面临的挑战和障碍：

1. 鼓励开展研究、制定指南，并确定相关资源，以支持在公共卫生领域生成人工智能应用的证据并扩大其规模。
2. 对实施人工智能和支持其应用所需的基础设施进行现代化改造。

下面，我们将讨论相关背景、HHS 迄今为止采取的行动，以及推动健康人工智能创新和应用的计划。

1. **鼓励开展研究、制定指南，并确定相关资源，以支持在公共卫生领域生成人工智能应用的证据并扩大其规模**

背景：

随着人工智能的不断进步，其全面影响仍不确定，这凸显了跨学科研究的必要性，以鼓励广泛创新、应用，并确保技术的规范使用。因此，针对人工智能在核心公共

卫生目标（如健康公平、患者隐私）以及不同公共卫生领域（如免疫推广、新兴疾病研究）的潜在影响开展有针对性的研究，能够为人工智能的有效性和跨领域适用性提供证据。HHS 及其下属部门可以通过确定并优先考虑可扩展的、具有高影响力的人工智能应用案例来发挥带头作用，这些案例应致力于解决最紧迫的公共卫生挑战，包括改善疾病监测和应急响应、解决资源有限和劳动力短缺问题，以及推进健康公平和医疗服务可及性等。HHS 将继续创建相关项目、指南和资源，以支持人工智能创新，并与更广泛的公共卫生生态系统分享研究成果，以鼓励进一步创新。

HHS 迄今为止采取的行动（非详尽列举）：

- 美国疾病控制与预防中心（CDC）的数据现代化倡议（DMI）正在对工具和技术（如先进的疾病监测系统、实时数据分析平台）进行投资，以便在公共卫生的各个层面获得更好、更快速且可行的见解，为决策提供支持（更多详细信息见上文）。
- CDC 的人工智能加速倡议（AIX）专注于将四个具有高影响力的公共卫生应用案例投入实际运作并扩大规模。
- CDC 员工聊天机器人是一个内部人工智能聊天机器人，用于提供与生成式人工智能交互的指南，使员工能够安全、规范地进行创新。
- 公共卫生数据战略人工智能计划里程碑 2.05 概述了该机构将如何利用人工智能并开展试点项目的计划。CDC 希望通过完成这一里程碑，鼓励安全、规范地使用人工智能，并提高公共卫生效率、应急准备能力和成果。
- 美国国立卫生研究院（NIH）的拨款和其他资源项目，如美国国家科学基金会（NSF）23 - 610：国家人工智能研究机构，或 NIH 的 Bridge2AI，为推进人工智能在生物医学和科学应用中的使用提供了资源。
- 美国食品药品监督管理局（FDA）在内部探索了人工智能的应用，包括但不限于：在 FDA 不良事件报告系统（FAERS）中对非公开不良事件数据进行去重；使用术语识别和新型合成阿片类药物检测与评估分析工具识别与阿片类药物相关的新术语，该工具利用公开可用的社交媒体和法医化学数据来识别药物产品在社交媒体文本中的新指代；以及使用基于人工智能的自然语言处理工具 ASSIST4 Tobacco 对烟草授权申请进行搜索和索引。

HHS 近期优先事项：

- 更新公共卫生数据战略，明确支持人工智能的开发和生命周期管理的措施。
- 确保 NIH 相关研究拨款继续允许使用人工智能，并研究其对公共卫生领域的影响。
- 继续开展人工智能试点项目，以加强对传染病爆发、慢性病和成瘾性物质的预测。
- 继续开展人工智能试点项目，针对提供者和患者，根据语言、文化水平和当地

情况，提供基于循证的公共卫生信息。

- 为公共卫生合作伙伴制定使用人工智能模型和公共卫生官员使用的人工智能系统的实施指南和操作手册，以支持其使用系统内常用工具进行现有业务。
- 与非营利组织及其他机构合作，在健康宣传活动中使用人工智能。

HHS 长期优先事项：

- 分享人工智能对公共卫生的影响及研究结果，包括运营影响、内部风险、益处及其他发现，为未来行动提供参考，并支持更广泛的社会共同体。
- 通过现有机制和适用的新机会，为公共卫生领域的人工智能应用提供资金和拨款支持。
- 考虑进行战略审查，并支持扩大与部门目标一致的高影响力投资，同时支持公共卫生合作伙伴在这些领域的协调一致。
- 考虑为其他利益相关者提供关于在何处以及如何扩大规模、何处可能存在投资机会的指南。

2. 对实施人工智能和支持其应用所需的基础设施进行现代化改造

背景：

许多公共卫生机构缺乏支持人工智能实施所需的现代技术基础设施。如前文所述，有效的公共卫生行动依赖于整合多种数据源（例如，通过公私数据共享以及将现有的个人健康数据与公共卫生数据相链接），以实现更全面的患者护理。然而，当前公共卫生数据系统相互孤立，现代化程度参差不齐，且通常依赖过时的技术运行，导致人工智能就绪程度各异。此外，数据格式的多样性和众多的数据标准限制了数据的和无缝共享。例如，美国疾病控制与预防中心（CDC）目前维护着 100 多个独立的疾病监测系统，但这些系统并未完全整合。由于这些技术和资源方面的挑战，公共卫生官员可能对使用人工智能解决方案持谨慎态度，而这些挑战影响了整个公共生态系统有效运行和沟通的能力。美国卫生与公众服务部（HHS），包括疾病控制与预防中心及其他部门，已开始为数据系统现代化奠定基础（例如数据现代化倡议），并且目前正在投入大量资源。然而，HHS 还可以且将继续采取更多行动。

HHS 迄今为止采取的行动（非详尽列举）：

- CDC 的数据现代化倡议（DMI）为州、地方、部落和领地（STLT）提供了直接资金和技术援助，以支持电子病例报告（自动数据传输）的实施、数据基础设施的现代化以及公共卫生数据系统的连接等工作。
- 截至 2024 年 9 月，公共卫生基础设施拨款已分配或发放了 6.11 亿美元资金，用于支持公共卫生数据现代化。这是拨给全国各地卫生部门 42 亿美元公共卫生基础设施拨款的一部分，旨在满足他们最紧迫的需求，涵盖从劳动力发展到实验室信息系统等领域。

- CDC 和助理部长办公室（ASTP）的联邦互操作倡议建立了可信交换框架和共同协定（TEFCA™），并使用了基于快速医疗互操作性资源（FHIR）的标准，在某些经过认证的健康信息技术应用和美国互操作性核心数据（USCDI +）数据项中来实施应用程序编程接口（API）。
- HHS 近期优先事项：
- 推进 HHS 数据战略，实现跨机构数据共享，以支持公共卫生领域的人工智能开发。
- 开展人工智能试点项目，协助整合和映射异质结构和非结构化公共卫生数据流以及与公共卫生相关的数据（例如，环境数据、社交媒体数据、零售数据和非处方药品销售数据）。
- 开展多辖区数据聚合试点项目，用于人工智能开发、验证和风险监测。
- 制定开发人工智能的战略，以推动将公共卫生功能整合到电子健康记录（EHR）系统中。
- 定期召开公共卫生合作伙伴论坛，就数据现代化工作进行协作。

HHS 长期优先事项：

- 根据可用资金和支持情况，为数据现代化和人工智能就绪计划提供更多的资助机会。继续在核心公共卫生数据系统中实施数据标准（例如，扩大 USCDI + 数据项的使用范围，并标准化常见数据指标/变量的定义，如人口数据），以提高数据的质量和完整性，并最大限度地提高人工智能的准确性和有效性。
- 继续当前的工作，简化技术环境，帮助公共卫生机构更好地整合和处理数据（例如，通过实施 CDC 关键的企业级综合数据平台，并帮助各辖区迁移并使用基于云的解决方案）。
- 继续推动互操作性标准的完善，使人工智能数据系统能够“相互理解”，包括标准化实施 TEFCA
- 继续为现有核心数据系统（如生命记录系统）的内部运营能力和数据现代化提供资金支持，以提高处理速度并获得更多洞察，例如识别与阿片类药物相关的死亡、药物过量及其他情况的趋势。
- 考虑更多整合公共卫生和医疗保健数据系统的方法，或提供“沙盒”试点机会。从战略上探索人工智能如何既能改进当前的现代化工作，以及当前的现代化工作如何可作为一个平台，鼓励其他机构使用人工智能工具。

5.6.2 促进可信赖的人工智能开发以及合乎道德和规范的使用

背景：

人工智能具有变革潜力，能够改变公众、患者和医疗服务提供者与医疗保健系统的互动

方式。这包括通过使用人工智能聊天机器人、人工智能翻译工具和其他服务，根据语言、地理位置和背景提供更个性化的健康信息。然而，要使这些技术发挥强大作用，利益相关者需要充分相信其能力，信任数据管理方式，并了解最佳使用实践。HHS 将继续致力于在确保安全和隐私的同时支持人工智能应用的创新。

HHS 可以在多个方面发挥重要作用，以实现人工智能的规范使用，包括：

1. 建立保障措施，确保数据质量和准确性。
2. 在公共卫生生态系统中统一数据安全政策。
3. 开发能够全面考虑和评估健康公平性的人工智能工具和技术。

下面，我们将讨论相关背景、HHS 迄今为止采取的行动，以及促进可信赖的人工智能开发以及合乎道德和规范使用的计划。

1. 建立保障措施，确保数据质量和准确性

背景：

人工智能模型可能面临数据质量和精度问题，在公共卫生领域尤其如此，不准确的数据可能危及个人和社区。无论是公共卫生机构开发的模型，还是缺乏公共卫生专业知识的机构（如一些科技公司）开发的模型，都必须使用适当的数据和参数进行训练，以确保其准确、可靠，并能够抵御滥用，从而获得广泛信任。所有模型都可以受益于强大的保障措施，以确保质量并控制共享信息，例如使用过滤器去除敏感内容并验证健康数据，以及进行持续监测以实时检测异常情况。组织可能会无意或恶意地创建有偏差的模型，或者使用错误数据传播错误信息和引发不信任，而这些负面结果很难识别和缓解。在国外，包括世界卫生组织（WHO）的报告和欧盟的《人工智能法案》在内的公共卫生机构最近发布了相关指南，以应对其中一些挑战。HHS 正在积极探索这一领域，并将继续制定机制、达成共识，并与其他机构合作支持建立和监督人工智能标准。

HHS 迄今为止采取的行动（非详尽列举）：

- 2024 年初，CDC 发布的《人工智能使用指南》为在公共卫生等领域中生成式人工智能的规范使用、开发和采购制定了原则和实践方法。
- 《HHS 在公共福利中规范使用人工智能计划》概述了 STLT 在管理公共福利（如健康筛查）时，在自动化和算法系统中规范使用人工智能的方式。
- 2023 年 6 月，NIH 院外研究办公室发布了 NOT - OD - 23 - 149《禁止在 NIH 同行评审过程中使用生成式人工智能技术》，禁止 NIH 科学同行评审人员使用自然语言处理、大语言模型或其他生成式人工智能技术来分析或撰写针对拨款申请和研发合同提案的同行评审意见。
- 2023 年 6 月发布的 CDC《发病率和死亡率周报》作者指南，为在医疗保健和公共卫生等领域的研究报告中使用人工智能提供了指导方针。

HHS 近期优先事项：

- 提高公共卫生领域数据和人工智能使用的透明度，以消除公众在关键领域的不信任，包括在疾病检测和监测中使用数据以及医疗错误信息和虚假信息的传播。
- 通过发布人工智能系统信息、模型卡片、培训信息和开源系统出版物，促进创新共享和最优方法的传播。
- 通过通俗易懂的公共卫生宣传和沟通工作，支持生成式人工智能的安全和规范使用，例如 CDC 的“Clean Slate”项目，该项目可以强调不当使用的风险并概述最优方法。
- 制定关于科学研究和公共卫生传播中人工智能系统作用的透明度标准和指南。
- 制定标准和指南，确保公共卫生提供者在使用人工智能时遵守现有的联邦民权法。

HHS 长期优先事项：

- 设计一种与人工智能系统设计者合作的机制，确保人工智能模型无法基于可能威胁公共卫生的医疗错误信息或虚假信息进行预训练。这包括确保人工智能系统的输出适当背景和信息，用于共享任何医疗相关信息。这可以通过与私营部门的人工智能培训组织合作来实现，以支持更广泛的应用。
- 继续与各组织合作，识别和缓解公共卫生领域的错误信息；在适当情况下支持合作性伙伴关系。
- 考虑实施对人工智能应用的持续监测和评估，以检测 and 解决潜在问题（例如，恶意行为者创建的模型），并在适当情况下与其他组织合作。
- 根据新的能力、联邦人工智能政策和 STLT 人工智能政策，更新和监测适用于整个公共卫生数据生态系统的现有公共卫生数据和人工智能治理结构及指南。

2. 在公共卫生生态系统中制定标准化的数据安全政策

背景：

许多现有的联邦和 STLT 层面的数据政策和指南在最初制定时并未考虑到人工智能技术。由于人工智能模型通常使用个人身份信息 (PII) 或受保护健康信息 (PHI) 数据进行训练或加权，在缺乏充分的数据保护和政策的情况下使用人工智能，可能会对患者隐私和安全构成重大风险。整个生态系统中缺乏统一的政策，还可能导致不同实体在管理和保护敏感数据方面存在不一致性，增加数据泄露的可能性，并可能导致声誉损失以及法律或财务后果。可以在 HHS 现有工作（如 HHS 网络安全计划或《健康保险流通与责任法案》(HIPAA) 安全规则）的基础上制定额外的指南，并针对公共卫生领域进行调整。

HHS 迄今为止采取的行动（非详尽列举）：

- HHS 通用数据使用协议 (DUA) 结构政策支持 HHS 在确保数据安全和伦理合规的前提下，向联邦机构或外部组织共享数据。

- HHS 医疗保健和公共卫生 (HPH) 网络安全目标包含了医疗保健组织和医疗服务提供组织的最优方法。

HHS 近期优先事项：

- 为公共卫生生态系统中的人工智能使用制定道德准则，以帮助保护个人权利和安全。
- 在公共卫生生态系统中推广关于安全开源软件和人工智能系统中数据安全实践的指南。

HHS 长期优先事项：

- 继续现有工作，对数据基础设施进行现代化改造，包括标准化核心公共卫生数据源，并通过安全措施（例如，实施隐私保护记录链接 (PPRL) 和 PII 减少技术，以防止敏感信息的共享）加强对个人数据的隐私保护。

3. 开发能够全面考虑和评估健康公平性的人工智能工具和技术

背景：

人工智能有潜力通过改善医疗服务提供、减少人为决策中的偏差，以及识别影响健康结果的可变根本“驱动因素”（例如，社区环境），而非仅仅依赖种族和性别等人口统计数据，来推进健康公平。然而，必须持续调查和解决人工智能可能无意中引入或放大健康差距的问题（例如，数据中的偏差可能导致算法出现偏差，对某些人群产生不成比例的影响）。特别是在服务不足的社区，由于先前的事件或不当的数据使用可能已经侵蚀了信任，人们可能对人工智能的开发和使用持怀疑态度。HHS 将继续努力促进以推进健康公平的方式使用人工智能。

HHS 迄今为止采取的行动（非详尽列举）：

- CDC 与佐治亚理工学院研究机构 (GTRI) 合作，召集 CDC 科学办公室和 GTRI 的专家，为公共卫生研究人员制定应对与人工智能使用相关的健康公平挑战的指南和培训资源。
- NIH 的人工智能促进健康公平与加速发现 (AIM - AHEAD) 计划旨在培养一支精通人工智能的多元化研究人员队伍，以满足代表性不足社区未满足的需求。
- 更多信息，请参见上文 CDC 的《人工智能使用指南》。

HHS 近期优先事项：

- 收集资源并举办一次公开教育活动，分享针对人工智能合成内容的潜在危害的缓解措施，这些内容往往旨在骗取高危人群的资源。
- 基于行业最优方法为公共卫生合作伙伴和外部合作伙伴制定模型卡片和系统卡片标准，用于记录人工智能系统，包括预期用途、已知限制、潜在模型偏差等关键领域最优方法。

HHS 长期优先事项：

- 与公共卫生不同领域的合作伙伴共同制定指南和最优方法，以保护健康、拯救生命，并减轻人工智能在各个领域造成的危害。
- 与相关实体协调，开展教育宣传活动，向高危人群宣传深度伪造和与人工智能相关的错误信息对公共卫生的潜在危害（例如，通过冒充医疗服务提供者泄露健康数据，传播看似来自可信公共卫生机构的虚假图像和视频）。
- 就人工智能开展公众教育和社区参与活动，这包括积极让家庭、社区和其他利益相关者参与公共卫生活动的开发和实施。这包括根据社区的需求水平提供资源，并建立一种双向关系，以建立信任、共享权力并进行合作，以支持所有相关方。

5.6.3 普及人工智能技术和资源**背景：**

人工智能为服务不足的人群以及资源匮乏的医疗保健系统和机构带来了巨大机遇，因为它有助于改善成本结构、解决资源和人员短缺问题，并优化整体资源分配和利用。与公众服务等领域相比，全球数据共享在公共卫生领域尤为重要。疾病无国界，只有通过透明的沟通与协作，才能迅速识别和控制疫情与病原体。公平获取人工智能技术能带来巨大益处和高投资回报，在多个领域扩大其影响力。HHS 可以通过以下方式应对当前的挑战：

1. 营造一个能够在公共卫生生态系统中实现数据共享的环境。
2. 支持人工智能的应用、开发与协作，尤其是为资源可能有限的 STLT 和社区组织提供支持。
3. 开发用户友好、可定制的开源人工智能工具，以扩大获取途径并满足不同用户的需求。

下面，我们将讨论相关背景、HHS 迄今为止采取的行动，以及普及人工智能技术和资源的计划。

1. 营造一个能够在公共卫生生态系统中实现数据共享的环境**背景：**

截至 2023 年，CDC 维护着一个高度复杂的数据基础设施，拥有超过 1000 个数据系统，这增加了能力现代化、实施人工智能基础设施以及确保最少数据录入等方面的挑战。这个数字还不包括由 STLT 机构拥有和运营的本地系统，而这些本地系统对于开展实地公共卫生活动和应对疫情至关重要。州和地方公共卫生官员收集和分析数据，根据这些数据向地方和州领导提出建议，并汇总这些数据以辅助联邦决策。有效的数据共享协议能够实现整个生态系统中从 STLT 和社区组织到联邦机构的快速、准确、双向数据共享，使各方能够可靠地了解美国不同地区当前的健康状况。

HHS 将继续支持有效的数据共享，以促进人工智能的应用。

HHS 迄今为止采取的行动（非详尽列举）：

- 2023 年，ASTP 发布了 TEFCA™，这是一个由国家卫生信息技术协调员办公室（ONC）管理的全国性健康数据交换框架，旨在帮助创建一个可靠的全国通用运营框架。全国有超过 50 个公共卫生辖区使用 TEFCA™ 交换来支持电子病例报告。
- 用于公共卫生的 USCDI + 是 CDC 和 ASTP 的合作项目，旨在在 USCDI 的基础上开发标准化的公共卫生数据项。
- ASTP 的 HTI - 2 拟议规则将使用公共卫生 API 和其他以公共卫生为重点的功能作为电子健康记录（EHR）可以获得认证的标准。此外，HTI - 2 拟议规则提议扩大 ONC 健康信息技术认证计划的认证标准，以包括适用于公共卫生信息技术系统的标准。

HHS 近期优先事项：

- 继续由联邦支持 TEFCA™ 健康数据交换框架，以简化医疗服务提供机构与公共卫生机构之间以及公共卫生机构之间的公共卫生信息共享。
- 继续推进用于公共卫生的 USCDI + 倡议，以提高全国公共卫生数据标准。

HHS 长期优先事项：

- 与标准制定组织就公共卫生领域人工智能技术的标准进行协调。

2. 支持人工智能的应用、开发与协作，尤其是为资源可能有限的 STLT 和社区组织提供支持

背景：

目前，创建人工智能工具需要大量的资金、数据和技术专长，所有这些都可能构成进入壁垒，使得人工智能供应商主要局限于私营部门或学术界。为应对 COVID - 19 大流行，联邦为数据现代化和支持系统提供的资金，使各组织能够开始更新数据系统、提高现有系统的效率并简化运营。然而，这些任务需要时间和大量投资，而这些资源在私营部门更容易获得。相比之下，公共卫生利益相关者，特别是 STLT 和非营利组织，将大部分资源用于维持基本运营和活动。他们可能在投资人工智能整合等长期需求方面能力有限，并且由于潜在的负面影响而风险承受能力较低。HHS 具备规模和能力，可通过提供资源、创建共享的集中式系统和标准，以及就如何鼓励创新和人工智能应用提供战略建议，来支持较小辖区和组织使用人工智能。

HHS 迄今为止采取的行动（非详尽列举）：

- 《HHS 在公共福利中规范使用人工智能计划》概述了对 STLT 在促进公共福利中使用人工智能方面的额外支持领域，包括向 STLT 提供可用资金信息。

- 更多信息，请参见上文 CDC 的数据现代化倡议（DMI）。

HHS 近期优先事项：

- 为地方组织开发具有人工智能增强功能的企业通信系统，以支持公共卫生宣传活动。
- 制定一项计划，利用 CDC 统一数据平台，为 STLT 提供工具、经过适当控制的数据、沙盒和基础设施，用于人工智能开发和实验。
- 与 STLT 召开公共卫生实践社区会议，以确定机会、发现推动因素和障碍，识别知识和资源共享的机会，并分享最优方法和经验教训（例如，通过专业协会）。
- 分享关于 STLT 和社区组织如何进行低成本、低风险“安全创新”的战术指南（例如，关于如何组建由现有员工组成的人工智能工作组，并测试利用现有或易于获取的技术和数据的简单人工智能应用案例的建议）。
- 鼓励并为 STLT 提供指南，使其尽可能使用现有的数据平台以及可用的人工智能系统和工具。

HHS 长期优先事项：

- 继续开展相关倡议，开发内部运营能力并对现有核心数据系统（如生命记录系统）进行现代化改造，包括开发和完善相关的人工智能基础设施能力。这项投资可以提高处理速度并提供洞察，例如识别与阿片类药物相关的死亡、药物过量的趋势和疾病路径。
- 支持提升生命统计运营链的系统能力，通过与开放文本字段和国际疾病分类第 11 次修订本（ICD - 11）的协作，增强洞察能力和自然语言处理能力。
- 根据可用资金和支持情况，为数据现代化和人工智能就绪计划提供更多的资助机会。
- 继续在核心公共卫生数据系统中实施数据标准，特别是在 STLT 和社区组织中（例如，扩大 USCDI + 的使用范围，并标准化常见数据指标/变量的定义，如人口数据）。
- 继续朝着整个生态系统的互操作性标准努力，使数据系统能够“相互理解”，包括人工智能的标准化实施。
- 实施一系列高价值、可扩展的人工智能项目，旨在改善公共卫生的特定领域。这些项目旨在提供此前没有人工智能技术就无法实现的即时、实际影响，助力解决现有和新出现的公共卫生问题（例如，使用人工智能从卫星图像中识别冷却塔，可以在军团病爆发期间更好地指导应对工作）。

3. 开发用户友好、可定制的开源人工智能工具，以扩大获取途径并满足不同用户的需求

背景：

在不同的公共卫生环境中，尤其是在资源匮乏的环境中使用人工智能，需要人工智能模型具有可定制性，并增加对高速互联网等技术和直观设计的人工智能工具的获取途径。增加低成本或无代码人工智能平台的可用性，并以低成本向公众提供，这能够使像 STLT 和社区组织这样的卫生机构开发出满足其社区独特需求的复杂模型。最近用于公共卫生领域人工智能的资源包括 ASTP 的健康信息技术领先边缘加速项目（LEAP），该项目提供资金以解决阻碍设计良好、可互操作的健康信息技术的开发、使用和/或推进的新出现的挑战。展望未来，HHS 可以考虑推进这些及其他工作，以支持开源人工智能工具的开发，特别是在这些工具可能产生最大影响且存在共享平台的领域。

HHS 迄今为止采取的行动（非详尽列举）：

- ASTP 的健康信息技术领先边缘加速项目（LEAP）为健康信息技术创新提供资金，以促进设计良好、可互操作的健康信息技术的开发、使用和/或推进。
- CDC 的人工智能加速倡议（AIX）正在开发具有高影响力的公共卫生人工智能试点项目，重点关注具有广泛可重用性且能解决常见公共卫生挑战的工具。

HHS 近期优先事项：

- 鼓励 STLT 利用通过地方政府项目提供的现有数据平台和开源人工智能系统。通过利用州数据平台获取人工智能资源，STLT 可以降低维护成本，并提升其合作伙伴（可能具有不同专业水平）的人工智能能力。

HHS 长期优先事项：

- 开发共享分析区域和工具，以促进联邦和 STLT 公共卫生合作伙伴对高价值人工智能应用案例的使用；确定常见的公共卫生挑战和数据平台，在这些地方使用这种方法可能会产生最大影响。
- 实施可扩展的由生成式人工智能驱动的聊天机器人，并使其易于 STLT 和其他公共卫生合作伙伴使用和修改。
- 开发并获取由人工智能驱动的开源工具以及配套的培训材料，以增强现有的公共卫生运营和工作人员能力。

5.6.4 培养具备人工智能能力的劳动力和组织文化**背景：**

公共卫生部门对于社区健康意识、预防和干预至关重要，但往往面临资源限制。如果在整合人工智能时能充分考虑社区需求，人工智能可以减轻公共卫生工作人员的负担。为了将人工智能整合到公共卫生运营中，营造学习和创新的环境，同时满足社区需求，HHS 可以支持开发应用案例、培训项目以及正式和非正式的人才培养渠道，使公共卫生工作者具备有效使用人工智能工具所需的技能。

HHS 可以通过以下方式应对当前的挑战：

1. 扩充和支持公共卫生劳动力，以解决人员倦怠和流失问题，同时提高效率和生产力。
2. 推广人工智能教育以及根据每个社区独特需求定制的基于社区的人工智能方法。

下面，我们将讨论相关背景、HHS 迄今为止采取的行动，以及培养具备人工智能能力的劳动力和组织文化的计划。

1. 扩充和支持公共卫生劳动力，以解决人员倦怠和流失问题，同时提高效率和生产力

背景：

如前所述，在 COVID - 19 大流行之前，公共卫生劳动力就面临挑战，而 COVID - 19 大流行加剧了劳动力问题，加速了人员倦怠和流失。对于使用人工智能，人们最大的担忧之一是它可能取代或减少现有工作岗位和员工数量；然而，公共卫生目前正面临严重的劳动力短缺。在 2017 年至 2021 年期间，近一半的州和地方公共卫生专业人员离职，如果这种流失率继续下去，公众可能无法应对未来的疫情和健康威胁。尽管人工智能无法取代对公共卫生知识共享和疾病应对至关重要的跨辖区和跨职能协作，但其具有巨大潜力，可以利用人工智能提高效率并支持人手不足的公共卫生劳动力队伍。例如，人工智能可以自动进行耗时或重复性的任务，使工作人员能够专注于更具战略性或以人为本的工作。同时，“人为干预”的方法可以确保在出现错误时进行监督和干预。HHS 可以继续识别和开发能够简化流程并提高现有公共卫生工作者生产力的人工智能应用案例。这不仅有助于缓解人员倦怠，还能促进整个公共卫生生态系统对人工智能的进一步理解和应用。

HHS 迄今为止采取的行动（非详尽列举）：

- 医疗保险和医疗补助服务中心（CMS）的人工智能手册包含教育材料，这些材料定义了医疗服务提供领域的人工智能应用案例和趋势，以及 CMS 目前正在使用和考虑在其自身运营中使用的应用程序，以及它们对患者护理的潜在影响（例如，可穿戴设备、数字孪生和客户支持）。

HHS 近期优先事项：

- 创建具有图像/音频编辑功能的生成式人工智能工具，以增强工作人员在教育 and 宣传方面的能力。

HHS 长期优先事项：

- 确定人工智能可以通过自动化报告和数据录入等常规和重复性任务来提高效率的机会。
- 投资培训和变革管理倡议，以提高人工智能的应用率，强调人工智能对提高劳动力效率和健康结果的影响，特别是在自动化常规和耗时任务方面。
- 全面评估人工智能对劳动力的潜在影响，以及公共卫生领域运营可能发生的变化（例如，对员工联系感和使命感的影响）。

2. 推广人工智能教育以及根据每个社区独特需求定制的基于社区的人工智能方法

背景：

基于社区和以人为本的方法在公共卫生领域广泛应用，从研究问题的选择到项目的实施，社区成员都参与其中，并被邀请利用他们的生活经验来识别和实施适当的干预措施。这些项目能够更好地解决导致健康问题的潜在风险因素，增强社区成员的能力并提高项目参与度，并且通常可以通过解决食物保障不足和交通不便等非医疗挑战的多因素方法来降低医疗成本。除了系统升级和资助项目外，还需要一支具备人工智能能力的劳动力队伍，他们了解何时以及如何使用人工智能（以及何时不使用），以及如何让社区参与进来，以确保人工智能得到规范和有效的使用。

HHS 迄今为止采取的行动（非详尽列举）：

- 参见上文关于 NIH 的人工智能促进健康公平与加速发现（AIM - AHEAD）计划的信息。

HHS 近期优先事项：

- 确定 HHS 在提高公众对公共卫生人工智能的认识和建立信任方面的战略重点。
- 与学术界和公共卫生学院协调，确保学生通过课程学习、学位项目和其他教育机会获得开展规范和符合道德的人工智能工作所需的技能。
- 与公共卫生合作组织和专业组织合作，将核心人工智能技能纳入沟通、能力培养和认证体系。
- 开发使用社区需求方法的人工智能项目和工具，在公共卫生项目的设计和实施过程中纳入社区的声音。

HHS 长期优先事项：

- 扩大现有的教育途径，为 STLT 和联邦工作人员提供提升运营人工智能和高级数据科学能力的机会。

5.7 结论

人工智能技术为加速公共卫生机构的运营效率、推进数据收集、预测和分析，以及改进宣传和沟通工作提供了独特机遇，且能以促进公平并改善健康结果的方式实现。然而，随着人工智能的广泛应用，也带来了诸多风险，例如可能通过深度伪造技术发起人工智能驱动的错误信息宣传活动，传播有害的健康建议。在未来几年，HHS 可以在应对 COVID - 19 大流行期间奠定的数据现代化和创新基础上，（1）推动对具有高影响力、可扩展的人工智能应用案例的投资和创新；（2）促进人工智能合乎道德、规范和可信赖的开发与使用；（3）普及人工智能技术和资源；（4）扩大劳动力的人工智能能力。通过与公共卫生生态系统中的利益相关者合作，HHS 能够朝着一个未来努力，在这个未来中，利用前沿技术，诸如由生成式人工智能驱动的聊天机器人分享基本健康信息，以及

通过深度学习基因组算法实现精准公共卫，能够帮助所有美国人实现最高水平的健康。HHS 致力于随着技术和应用案例的不断变化而不断发展其人工智能战略，以更好地改善公众健康。

6 网络安全和关键基础设施保护

6.1 引言和背景

保护数字系统免受网络威胁对于实现新兴技术（如人工智能）的益处并将其风险降至最低至关重要。如果没有有效的风险管理，人工智能系统可能会危及患者、参与者和公众的安全，泄露个人身份信息，并削弱公众对医疗保健和公共卫生系统的信任。然而，通过适当的控制措施，人工智能对美国健康和公众服务生态系统的潜在益处是巨大的。此外，应对网络安全风险对于遵守第 14410 号行政命令《安全、可靠和可信赖的人工智能开发与使用》至关重要，该行政命令要求 HHS 和联邦政府在关键基础设施领域促进人工智能模型的安全、可靠设计、开发和部署。为响应该行政命令和《国家网络安全战略》，HHS 于 2023 年 12 月发布了其《网络安全战略》，概述了改善健康和公众服务领域网络安全的行动。本文件在 HHS《网络安全战略》的基础上，重点介绍自该战略发布以来该部门采取的新行动，并概述其他优先事项。

网络事件对美国医疗保健系统的威胁是真实且不断增长的。医疗保健占美国国内生产总值的 4.5 万亿美元（17%），占美国就业的 9%。这些因素使医疗保健成为一个庞大的攻击目标。根据一项调查，92% 的医疗保健组织在过去 12 个月中至少经历了一次网络事件，HHS 民权办公室 (OCR) 报告称，涉及勒索软件的大规模数据泄露事件从 2018 年到 2022 年增加了 264%。

在健康和公众服务领域，网络安全事件会影响多个利益相关者。以往发生的意外导致了患者护理的延误，以及医疗服务提供者、支付方和州公共卫生部门在运营和财务上的中断。此外，人工智能的引入扩大了威胁范围，因为人工智能应用越来越多地被网络攻击者用作工具，成为数字系统中可被利用的漏洞，但同时也可作为新的防御工具。随着人工智能在医疗保健和公共卫生生态系统中的应用规模不断扩大，网络安全保护措施也必须相应扩展。

在本章中，HHS 概述了网络安全风险的当前和预期趋势，人工智能如何影响并造成这些风险，以及它们对医疗保健、公共卫生和公众服务的影响。然后，该部门概述了采取行动以更好地使各组织应对这些威胁的机会、HHS 正在采取的行动，以及可进一步增强健康和公众服务生态系统网络安全能力的其他行动。

6.1.1 行动计划摘要

在本章后面部分，HHS 阐述了为提高该部门管理其网络安全要求的能力而提议采取的行动。以下是这些行动的主要主题。有关提议行动的详细信息，请参见 6.4 行动计划。

行动主题：

1. 解决健康和公众服务领域中具备适当技能的网络安全工作人员短缺问题，以填补相关岗位空缺。
2. 支持最优方法的标准化和协调统一，特别是在网络安全治理方面。

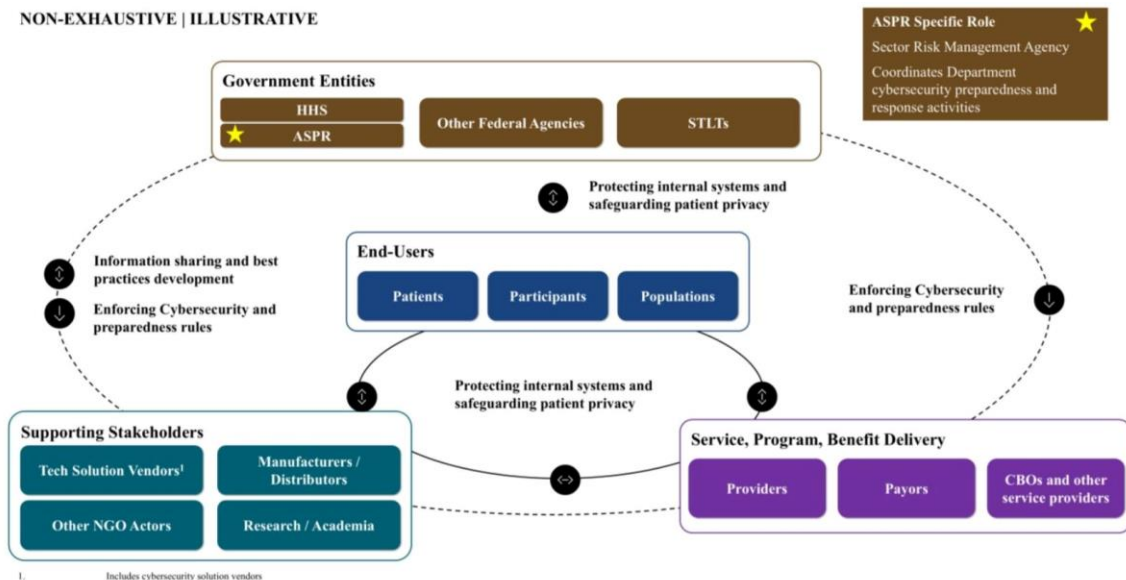
3. 减少并管理实施新的网络安全能力时的复杂性。
4. 明确在健康和公众服务领域中，应对隐私与公平以及隐私与安全之间尖锐矛盾的方法。

6.2 参与健康和公众服务生态系统中网络安全和关键基础设施保护的利益相关者

HHS 在促进网络安全方面发挥着双重作用：首先，通过信息共享和最优方法开发，作为该领域的合作伙伴；其次，作为监管机构，执行网络安全和应急准备规则。除 HHS 外，联邦政府其他部门、STLT、医疗服务提供者、支付方、社区组织和其他非政府利益相关者都有责任抵御网络威胁，并维护其组织的网络安全能力。

HHS 内的多个部门和团体在网络安全方面发挥着作用。其中包括卫生部门协调委员会（HSCC）和 HHS 405(d)工作组，这两个公私合作伙伴关系有助于为医疗保健、公共卫生和公众服务领域开发和共享人工智能指南。助理部长办公室负责准备和响应（ASPR）代表 HHS 协调医疗保健和公共卫生部门的部门风险管理机构活动，协调 HHS 内部、联邦机构之间以及与行业合作伙伴的网络安全应急准备和响应活动。

图表 15 展示了参与网络安全和关键基础设施保护的利益相关者之间的示例交互流程，该图表并非详尽无遗，仅作说明之用。请注意，该图表并未涵盖所有利益相关者的角色和互动。有关监管指南和权限的更多详细信息，请参考其他 HHS 文件。角色可能因医疗保健、公共卫生或公众服务生态系统的领域或部分而异。



6.3 网络安全和关键基础设施保护的趋势

1. **医疗保健行业和全球范围内的网络事件呈上升趋势，与网络犯罪相关的成本也在不断增长：**随着公司、机构和组织的转型和现代化，网络威胁的数量和类型每年都在

增加。一项分析估计，2023 年至 2024 年期间，各行业的安全漏洞将增加 589%。此外，2023 年网络事件的成本估计达到 8 万亿美元，并且还在持续增长。健康和公众服务组织也面临着日益增加的网络安全威胁，包括勒索软件攻击、网络钓鱼攻击、第三方数据泄露、数据泄露和社会工程攻击。医疗保健领域的数字化加速（例如电子健康记录的使用）使医疗保健成为网络威胁的高优先级目标，并加剧了建立有效防御措施的复杂性。2021 年至 2022 年期间，卫生和公共卫生部门的勒索软件事件增加了 42%，自 2016 年以来，影响卫生系统的网络事件发生频率翻了一番。这些事件可能导致系统故障，危及患者安全，并产生其他后果。

2. **网络犯罪行业规模庞大且成熟，有能力对健康和公众服务组织发动日益复杂的攻击：**网络犯罪是一个价值数十亿美元的复杂行业，它拥有研发功能，且攻击能力不断提升。攻击者正在使用包括人工智能在内的新工具，将端到端攻击生命周期从数周缩短至数天甚至数小时。近年来，攻击者利用公共卫生危机展示其攻击手段的威力。例如，在 COVID - 19 大流行期间，由于威胁活动增加以及向居家办公模式转变导致的脆弱性增加，勒索软件和网络钓鱼攻击在全球范围内激增。
3. **人工智能（尤其是生成式人工智能）的使用预计将增加网络威胁、漏洞以及错误和事故的可能性：**美国联邦调查局警告称，网络犯罪分子越来越频繁地利用人工智能工具策划有针对性的网络钓鱼活动。例如，使用大语言模型模拟人类的人工智能驱动的社会工程攻击正变得越来越成功。仅在 2023 年前两个月，新型网络钓鱼攻击就激增了 135%。此外，新的恶意软件不断涌现，其能够规避端点检测与响应（EDR）技术等传统网络安全工具。对于医疗保健组织而言，人工智能驱动的网络钓鱼攻击是美国医疗保健网络威胁中最常用的攻击手段之一。在公共卫生领域，人工智能生成的深度伪造技术可用于传播错误信息，这可能会降低人们寻求治疗的意愿，或者仅仅是破坏对公共卫生机构的信任。此外，人工智能驱动的系统还可用于对敏感健康信息进行去匿名化处理，从而引发代价高昂的勒索软件攻击。考虑到医疗保健和公共卫生领域有大量用于临床试验、精准医疗和医学研究的匿名数据集，这尤其令人担忧。其他特定于医疗保健的威胁可能包括对医学成像的对抗性攻击或数据投毒等手段，而影响联邦机构或 STLT 的威胁包括自动化社会工程攻击或虚假信息宣传活动。还有一系列其他对抗性人工智能技术处于不同的开发阶段，复杂程度各异。
4. **健康和公众服务领域对大型数据集的需求和访问不断增加，也导致数据泄露风险加剧：**尽管医疗保健在使用云存储方面落后于其他行业，但在线患者平台的增加、人工智能的应用以及电子健康记录的使用，导致更多的健康数据存储在云端。医疗保健、公共卫生和公众服务组织管理着大量敏感数据集，包括受保护的健康信息，并且许多利益相关者都可以访问这些数据。随着数据越来越多地迁移到云存储，所有组织都必须采取网络安全预防措施来保护敏感数据。此外，依赖第三方云存储可能会放大漏洞。事实上，35%的医疗保健数据泄露涉及第三方供应商。医疗保健数据泄露的后果是巨大的。例如，2 月份的一次勒索软件攻击暴露了 1 亿人的私人健康信息，可能造成超过 25 亿美元的财务影响。医疗保健数据泄露的成本越来越高，这是由于业务中断、客户支持和修复所带来的损失。现在，一个组织遭受医疗保健数据泄露的平均成本为 1000 万美元。

5. **传统的网络威胁应对工具仍然有效，但由于多种挑战，组织的网络风险应对能力滞后：**尽管人工智能驱动的网络威胁可能更具破坏性，但它们的攻击途径往往与传统网络安全攻击相似。最近的数据表明，84%的关键基础设施事件涉及“通过资产和补丁管理、凭证强化以及最小权限原则等最优方法和安全基本措施本可以缓解的初始访问途径”。传统的网络安全实践仍然有助于抵御这些威胁。然而，由于网络安全劳动力技能不匹配、最优方法缺乏标准化、实施复杂性以及其他障碍等挑战，组织在使用传统的网络威胁应对工具时也面临困难。在下一节中，HHS 将提供这些挑战的更多背景信息，并概述该部门可采取的行动机会，以增强该领域的网络安全和关键基础设施保护能力。

6.4 行动计划

健康和公众服务组织正在加大对其网络安全能力的投资。一项估计表明，2020 年至 2025 年期间，全球医疗保健行业将在网络产品和服务上花费 1250 亿美元，年增长率为 15%。2023 年对医疗保健网络安全专业人员的一项调查发现，超过一半的人在过去一年中看到其网络安全预算有所增加。尽管关注度和投资有所增加，但医疗保健组织仍难以应对不断升级的威胁。例如，2022 年至 2023 年期间，医疗保健行业的勒索软件攻击几乎翻了一番。此外，网络安全支出的重点也发生了变化；从 2016 年到 2022 年，用于预防事件的支出比例从 60% 下降到 30%，现在更多资源被分配用于管理正在发生的事件。

下面，HHS 概述了几个提高该领域管理其网络安全要求能力的行动机会。这些机会包括：

1. 解决健康和公众服务领域中具备适当技能的网络安全工作人员短缺问题，以填补相关岗位空缺。
2. 支持最优方法的标准化和协调统一，特别是在网络安全治理方面。
3. 简化并优化部署新网络安全方案时的复杂流程。
4. 在健康和公众服务领域中，明确解决隐私与公平以及隐私与安全之间尖锐矛盾的方法。

对于这些机会中的每一个，HHS 都提供了背景信息，并强调了其已经采取的缓解措施以及正在考虑的未来行动。

1. **解决健康和公众服务领域中具备适当技能的网络安全工作人员短缺问题，以填补相关岗位空缺**

背景：

许多组织缺乏抵御最新威胁所需的网络安全人才、知识和专业技能，并且在填补关键岗位方面面临困难。在美国，熟练网络安全工作人员的缺口扩大速度超过了新招聘的速度。这种短缺可能源于技能不匹配，而非缺乏求职的网络专业人员。更多传统的网络专业人员在云服务、人工智能和生成式人工智能数据与分析，或健康和公众服务信息技术等领域缺乏所需的专业知识。越来越多的非网络安全团队的领导者

也在寻找网络安全人才。然而，这些领导者往往缺乏评估候选人或有效满足其招聘需求所需的基本网络安全知识。在医疗保健领域，CDW 的一项调查显示，只有 14% 的医疗保健信息技术领导者表示其安全团队人员配备齐全。HHS 已经采取行动提高网络安全劳动力能力，如下所述，HHS 将寻求开发培训项目并探索资源配置，以吸引更多具备适当技能的人才进入该领域。

HHS 迄今为止采取的行动（非详尽列举）：

- 通过主动监测、数据共享和协作，提高资源匮乏的 STLT 的能力。HHS 继续监测和共享数据，包括人工智能威胁相关数据，以提高资源匮乏的 STLT 的能力，并与政府合作伙伴合作制定和共享保护人工智能模型的基本网络安全实践指南草案，继续提供工具和资源，帮助资源匮乏的实体实施强大的网络安全措施。
- 高级研究计划局卫生（ARPA - H）正在开发新的工具，能够自动检测和修复网络漏洞，减轻医院和医疗保健组织的网络安全负担。这些措施包括：
- 与国防高级研究计划局（DARPA）合作发起人工智能网络挑战赛，利用人工智能创建可用于识别和修复漏洞的自动工具，这些工具可部署在国家开源软件供应链中。
- 创建“自主防御通用补丁和修复”计划，旨在开发一种自主网络威胁解决方案，实现主动、可扩展和同步的安全更新，减少保障医院安全所需的不确定性和人工工作量。

HHS 近期优先事项：

- 针对组织领导和非网络团队的招聘经理，开发更多针对健康和公众服务领域的网络安全培训。
- 评估支持资源匮乏的医疗保健和公共卫生组织开展网络安全劳动力发展的机会。
- 支持 HHS、STLT 和社区组织使用支持安全数据共享的技术。

HHS 长期优先事项：

- 将新的网络安全要求纳入 HHS 的拨款、合同和合作协议中。
- 探索将人工智能驱动的威胁纳入 HHS 优先情报需求，以增加 HHS 内部以及医疗保健、公共卫生和公众服务领域之间现有的网络威胁情报共享。

2. 支持最优方法的标准化和协调统一，特别是在网络安全治理方面

背景：

网络安全涵盖了从战略规划到数据保护，再到应对攻击并恢复正常的一系列复杂能力。每个组织对其网络安全能力的重视程度和优先级各不相同。因此，对于何时以及如何使用可信架构技术，目前尚无公认的标准。在卫生领域，这些挑战可能延伸

到与网络相关的风险管理，例如有时在解决人工智能系统故障方面，角色和职责定义不明确，在使用人工智能系统进行决策时对责任的理解不足，以及无法验证模型输出的准确性。在许多组织中，这可能导致网络管理使用一种临时应对方式。通常负责信息技术相关威胁的网络安全团队（通常是首席信息安全官负责处理与信息技术相关的威胁），无法始终获得有效保护其组织所需的资源。此外，他们可能未能尽早参与评估新项目的网络风险，或者可能仅在发生数据泄露时才介入。这可能导致网络能力的差距，包括资产管理、漏洞管理、关键指标和报告、身份和访问管理以及数据保护等方面。HHS 已经采取行动使用并发布最优方法标准，并将随着网络安全实践的发展继续制定和更新指南。

HHS 迄今为止采取的行动（非详尽列举）：

- 使用了美国国家标准与技术研究院（NIST）的人工智能风险管理框架：将人工智能风险管理实践整合到网络安全、应急管理、临床运营、医疗设备、法律、劳动力管理、供应链和采购等方面的规划中。
- 发布了 HHS《网络安全战略》（2023 年 12 月），该战略建议实施基本的传统网络安全措施，旨在帮助所有组织通过适当的工具和措施提升其安全底线，以管控组织内人工智能带来的风险。
- 发布了其医疗保健和公共卫生（HPH）网络安全绩效目标（CPGs），帮助组织确定实施高影响力网络安全实践的优先级，这些目标是根据美国网络安全与基础设施安全局（CISA）的绩效目标以及行业最优方法改编而来，以适应医疗保健领域的背景。
- 发布了根据《健康保险流通与责任法案》（HIPAA）加强医疗保健网络安全的拟议措施（2024 年 12 月），要求健康计划、医疗保健结算所、大多数医疗服务提供者及其业务合作伙伴更好地保护个人的电子受保护健康信息，抵御外部和内部威胁。
- 与政府合作伙伴合作，制定和共享保护人工智能模型的基本网络安全实践指南草案。
- HHS 近期优先事项：
- 制定关于在部署人工智能的系统遭到破坏后维持运营的指南。用户应制定政策、配备工具并接受培训，以了解人工智能系统何时产生错误输出，并制定人工智能系统遭到破坏时的恢复计划。
- 更新现有的使用法规和指南，纳入维护网络安全的最优方法，包括维护安全的数据传输和共享方式。

3. 减少并管理实施新的网络安全能力时的复杂性

背景：

上述威胁表明，在人工智能应用开发的各个阶段，都需要强化数据安全和访问控制。

安全的设计和培训有助于防止人工智能产生虚构内容、数据泄露和数据窃取，这对于从事生物技术、新病原体等敏感研究的机构尤为关键。人工智能模型开发者必须在其运营的各个环节实施强有力的安全控制，而这些解决方案的用户（如医疗保健利益相关者）需要接受培训，以了解并安全地整合这些解决方案。然而，由于人员持续短缺、缺乏标准化以及技术环境日益复杂，组织在与解决方案提供商协调、改变组织规范以适应新部署，或者在整个企业范围内部署和扩展解决方案时困难重重。HHS 将通过以下潜在行动，如完善 HHS 医疗保健信息技术认证，寻找降低复杂性的方法。

HHS 迄今为止采取的行动（非详尽列举）：

- 发布增强软件透明度的指南。《联邦食品、药品和化妆品法案》第 524B (b)(3) 条要求“网络设备”的医疗设备制造商在上市前提交文件时，提供一种名为“软件物料清单” (SBOM) 的软件透明度机制。SBOM 是 FDA 评估设备安全性、上市后漏洞以及事故响应的一部分。FDA 将持续监测设备人工智能在上市前提交文件和上市后问题中的网络安全情况，以评估是否需要额外政策来保障患者安全。
- 数字健康安全 (DIGIHEALS) 计划正与人工智能和网络安全专家合作，通过改造为国家安全开发的成熟技术，使其能够应用于民用卫生系统、临床护理机构甚至个人健康设备，从而强化电子健康生态系统。

HHS 近期优先事项：

- 鼓励医疗保健信息技术开发者在产品中实施隐私和安全设计，包括在产品中内置网络控制功能，或提供服务应用程序编程接口，以便将网络控制功能集成到其他系统中。可通过以下方式实现：
- 强化 ASTP 的 ONC 医疗保健信息技术认证计划中的网络安全认证标准。
- 扩大 ASTP 的 ONC 医疗保健信息技术认证计划的范围，将更多医疗保健信息技术系统纳入其中（例如实验室系统、远程医疗——患者健康记录、交换或访问系统、临床医生主导的临床数据登记系统、电子事先授权系统和结算所处理系统）。
- 与医疗保健利益相关者合作，为希望评估人工智能对其组织构成风险的机构制定指南和资源。这可能包括收购和采购指南，以帮助小型 / 资源匮乏的组织评估不同人工智能工具和解决方案的安全影响，培训员工掌握最优方法，或评估被认为安全的数据传输和数据共享工具的风险。
- 绘制 HHS 人工智能系统价值链中的安全和隐私风险图，以解决第三方相关问题，包括受损的库、未经审查的数据或标签错误等。
- 为研究网络安全对临床环境的影响提供资金支持。

HHS 长期优先事项：

- 考虑与行业和其他政府机构合作，开发关键风险指标和性能阈值，以增强软件透明度。

4. 明确在健康和公众服务领域处理隐私与公平、隐私与安全之间尖锐矛盾的方法

背景：

健康和公众服务领域的网络安全团队需要在隐私和数据保护的需求与有时更广泛却相互矛盾的目标之间寻求平衡。例如，一个组织可能希望实施严格的数据隐私和数据共享限制，同时又旨在在人工智能模型中广泛纳入数据，以减少潜在偏差或监测人群健康状况。在实践中，如果没有标准或指南，这些相互冲突的优先事项可能导致决策实施的延迟，或选择出不够理想的设计，既无法充分保护隐私，也无法实现更广泛的健康相关目标。HHS 将努力阐明并界定网络安全方面的权衡，以协助利益相关者做出决策。

HHS 近期优先事项：

- 为组织提供指南，说明如何处理网络安全方面的权衡问题，以及在保护网络安全和隐私方面应重点关注的方向。

6.5 结论

在医疗保健、公共卫生和公众服务领域，网络威胁引发的干扰会直接影响人们的生活。HHS 及其更广泛的生态系统必须认识到网络威胁的增长，并采取措施确保其组织得到保护。对于希望以规范且有效的方式扩展人工智能应用的 HHS 生态系统中的任何组织而言，网络安全是一项基本能力。然而，如上文所述，尽管人们普遍认识到网络威胁，且对网络安全的关注度有所提高，但各组织仍难以跟上潜在攻击者的步伐。在提升网络安全人才技能、建立和推广最优方法及治理标准、降低引入新技术能力的复杂性，以及协助组织平衡网络安全和隐私问题等方面，仍存在巨大改进空间。

HHS 的网络安全和基础设施保护战略将随着威胁形势的变化而持续发展。HHS 致力于协助其下属机构以及更广泛的 HHS 利益相关者提升网络安全能力，并已为此采取了多项措施。该部门将继续考虑采取其他行动，提升医疗保健系统的安全底线，使各组织更轻松、安全地使用对美国民众产生积极影响的人工智能应用。

7 内部运营

7.1 引言和背景

人工智能为 HHS 带来了广泛的机遇。HHS 的运营和职能部门已经在利用人工智能推进其使命，以改善内部运营并提升面向公众服务的执行效果。HHS 对人工智能的广泛应用需要一个正式的、全部门统一的管理方法。该部门在人工智能方面的举措还必须聚焦于变革管理和适应性，因为人工智能的实施和应用可能会改变现有的流程。通过优化部门在采购、测试、部署和安全管理内部人工智能解决方案方面的流程、政策和架构，HHS 旨在加速知识共享，并协调对人工智能投资的支持。这将确保部门内部的一致性，同时也为各机构提供适当的灵活性以推动创新。

为与第 13859 号行政命令、第 14110 号行政命令以及管理和预算办公室（OMB）的 M - 24 - 10 和 M - 24 - 18 备忘录保持一致，HHS 的首席人工智能官办公室（OCAIO）将引领三个关键领域的工作，以便在 HHS 部门层面和各部门内部部署高价值、可信赖的人工智能：

1. 治理
2. 内部流程改进与创新
3. 劳动力与人才管理

为制定一个统一的战略，这些关键领域必须融入 HHS 各部门的主要内部运营中，并在部门各个层级得以实施。这种方法有助于在 OCAIO 的集中协调与 HHS 各部门为实现各自使命目标所需的灵活性之间找到恰当的平衡。此外，HHS 的人工智能战略将与现有的信息技术基础设施审查和部署政策、框架及法定职责保持一致。

7.2 机遇与风险

机遇：

在 HHS 内部运营中增加人工智能的使用和应用，带来了诸多重要机遇，包括：

1. 提升面向公众项目和服务的质量、体验及安全性：通过部署在私营或公共部门中已经过充分验证的应用案例，人工智能能够助力 HHS 每年更有效地为数亿民众提供健康和公众服务。例如，直接参与患者护理的 HHS 机构可以利用相关技术实现更精准的患者监测，而负责提供公众服务的机构则可以借助人工智能更高效地为受益者匹配最合适的服务。
2. 为 HHS 内部人工智能的创新应用提供政策、指南和流程支持：HHS 需要紧跟快速发展的技术生态系统，以顺利履行其使命。通过试点和部署有助于制定有效指南和改进流程的应用案例，HHS 能够更高效地运作，进而更好地为美国的健康和公众服务提供监督和服务。

3. 积累知识和能力，为 HHS 各领域的公共政策和指南提供参考：HHS 内部对人工智能的应用将增长部门在人工智能方面的知识和能力。这反过来又使 HHS 能够为这些技术在其负责的领域（如医疗服务提供和研发）提供更具针对性的政策、指南和监督。
4. 提高劳动力效率：人工智能有潜力将目前需要人工和承包商直接投入时间的手动流程自动化。利用人工智能辅助那些可通过技术优化的任务，能够让 HHS 的工作人员有更多时间投入到高影响力的活动中（例如审查、协调、执法以及在适用情况下直接提供护理服务）。

风险：利用人工智能支持 HHS 的使命，在增加一些现有风险的同时，也带来了新的风险。这些内部风险的示例包括：

- **数据隐私和安全风险：**随着 HHS 各部门在使用人工智能解决方案方面积累经验，未来的应用案例可能会使用涉及患者、参与者或社区层面隐私的信息来训练内部模型或生成输出结果。对于这类应用，需要像对待非人工智能应用案例一样保持高度警惕，并且根据具体情况和数据来源，可能需要额外的控制措施或使用其他技术，以确保人工智能的应用不会产生新的漏洞。
- **执行风险：**对人工智能的使用制定过于严格的内部指导方针，可能会抑制那些能够产生积极影响的创新方法。研究表明，对人工智能及其影响（例如对职业或项目参与者的影响）的先入之见，可能会阻碍新的人工智能技术在工作场所的成功部署。如果 HHS 与员工之间就人工智能的进展和挑战沟通不畅，可能会进一步加剧这种风险。
- **对劳动力培训和技能的影响：**将人工智能融入 HHS 各部门的技术和 workflows，可能会带来风险，例如如果员工不再从事曾经属于其职责范围内的任务，可能会出现技能差距。对于诸如计算、基本算术、日程安排等常规任务而言，这可能并非风险，但对于客户支持或其他需要人际互动的技能而言，可能会构成挑战。此外，对人工智能的使用可能会导致对其过度依赖，负责监督由人工智能支持的职能的工作人员可能无法进行充分的监督。

在下文“计划开展的 HHS 活动”中提出的行动建议里，将充分考虑这些风险。如果不能妥善管理，成功执行本战略规划还可能面临其他风险。在应对人工智能开发和/或使用所带来的这些或其他风险时，HHS 还将根据特定模型、工具或应用案例的预期风险水平，量身定制风险管理策略。包含人工智能的系统应该应用风险管理方法，在相关流程和系统框架的各个阶段识别、解决和监测潜在的负面影响。

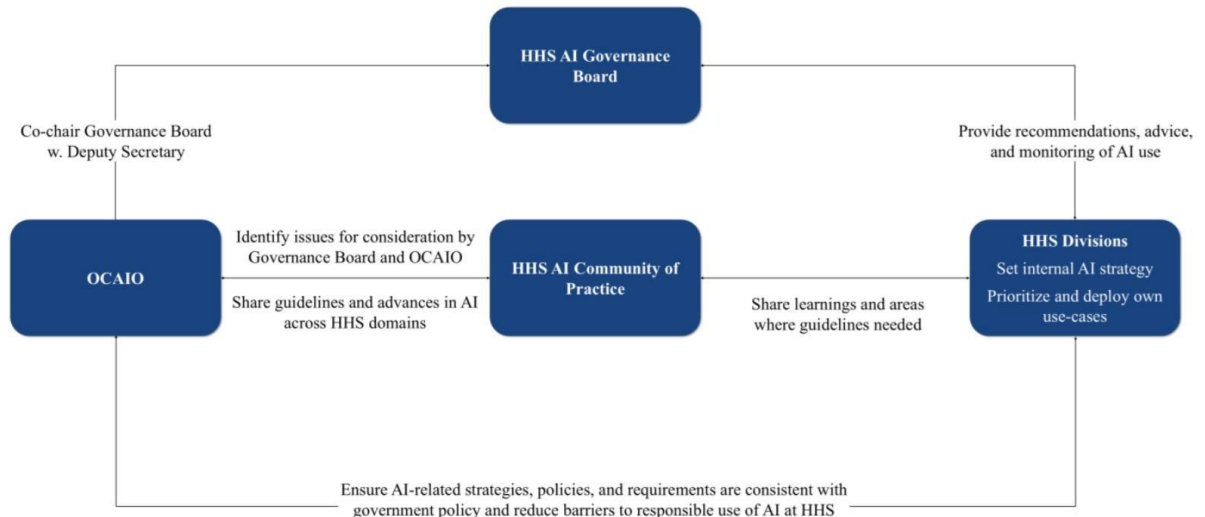
7.3 治理

背景：

在人工智能解决方案从概念提出到执行再到退役的整个生命周期中，有效的治理对于促进合理应用和风险管理至关重要。正如 HHS 及其各部门围绕信息技术构建了大量基础

设施，以确保规范地使用并最大程度减少因不当数据共享带来的风险一样，HHS 已经并将继续围绕人工智能建立必要的保障措施。人工智能治理实践将尽可能利用 HHS 现有的和/或部门层面的管理机构和流程，以确保战略一致性，并避免给 HHS 各部门带来不必要的负担。人工智能治理将根据每个部门的独特结构和需求采取量身定制的方法，在最大程度减少人工智能相关风险潜在影响的同时，促进创新。

图表 16 展示了 OCAIO 与 HHS 各机构之间的互动关系，并从宏观层面定义了 OCAIO 在 HHS 中的角色。图表中所示的治理机制将在下文进一步详细说明。



HHS 迄今为止采取的行动（非详尽列举）：

HHS 已经建立了基础的治理架构以支持人工智能的应用，包括：

- 根据 M - 24 - 10 的要求，聘请了常任首席人工智能官（CAIO）。OCAIO 将确保所有战略和部门政策、要求及指南与政府政策保持一致，并减少在 HHS 内规范地使用人工智能的障碍。虽然 OCAIO 对 HHS 内人工智能解决方案的治理负有主要责任，但它将与部门内的其他办公室协商合作，以确保这些政策和指南适用于 HHS 的各个部门。
- 成立了 HHS 人工智能治理委员会，作为主要的治理机构，负责指导 HHS 的人工智能政策、项目和技术应用，并确保这些政策符合 FAVES 原则。该委员会就人工智能使用的关键问题提供建议、指导并进行监督。委员会于 2024 年 5 月首次召开会议，由副部长担任主席，CAIO 担任联合主席，成员包括 HHS 各部门的高级领导。该委员会负责支持人工智能治理工作，制定全部部门的人工智能战略重点，并监督战略的执行情况。委员会还将监测 HHS 实施本战略计划的进展。
- 成立了由 OCAIO 管理的 HHS 人工智能实践团队（CoP），成员包括来自 HHS 各部门对人工智能感兴趣的工作人员。该实践团队的目标是为部门内的持续学习和协作提供机会，为 OCAIO 确定需要 HHS 层面和跨机构协调的优先事项，并帮助识别需要 HHS 人工智能治理委员会审议的关键问题。HHS AI CoP 还为人工智能政策制定、人工智能人才和劳动力发展等关键领域的工作组提供支持。

HHS 近期优先事项：

虽然 HHS 已经制定了上述治理方法，但要完善并实施更全面的架构和流程，以加速应用案例的使用，并确保组织为人工智能驱动的使命提升做好准备，仍需要时间。为进一步完善这些治理实践，HHS 的 OCAIO 将采取以下行动：

- 强化并正式建立全面的治理架构：HHS 将在上述基础要素的基础上进行扩展，以支持组织内人工智能的规范使用。这将包括明确阐述并记录关键治理机构之间的角色、职责和决策权。
- 为 HHS 各部门提供人工智能治理指南：这些指南将支持各部门制定执行其独特使命所需的特定治理政策，并将考虑各部门已建立的现有架构。
- 通过管理人工智能工作组丰富实践社区：这些工作组将分享部门内人工智能应用的实时见解，包括共享学习经验、最优方法，以及确认新出现问题的其他途径。HHS 还将继续探索随着时间推移扩展实践社区的方法，以满足部门的需求。
- 建立定期审查和修订人工智能治理架构的机制：HHS 人工智能治理委员会将制定定期审查 HHS 人工智能架构和指南的流程（例如每年审查一次）。

7.4 内部流程改进与创新**背景：**

HHS 必须确保其流程的设置有助于安全有效地使用和开发人工智能。这涵盖多种工作流程，包括采购（例如采购人工智能解决方案、在选择定制工具时使用人工智能）、原型制作、试点项目、部署（例如创建用于疾病流行监测的分析引擎）、维护和运营（例如确保持续的质量和合规性）以及安全（例如避免敏感数据的滥用）。同样，HHS 必须根据需要调整其在拨款制定、拨款监督和项目评估方面的内部流程，以与人工智能应用的最优方法保持一致，并维持项目性和科学性的完整及可持续。HHS 已经有多项政策指导这些领域的工作，人工智能的应用需要与现有方法保持一致。例如，无论是否使用人工智能，敏感数据存储都必须达到同样高的标准。在制定方法时，HHS 将评估是否需要更新现有政策（例如授权运营框架），以确保它们支持人工智能的使用和开发。这些政策还将遵循 OMB 的 M - 24 - 10 和 M - 24 - 18 备忘录，并将包括相关的额外步骤，以识别和缓解风险，例如当人工智能解决方案被认定为“影响权利”或“影响安全”时。

HHS 迄今为止采取的行动（非详尽列举）：

HHS 已经为人工智能的使用奠定了基础，包括：

- 根据第 13960 号行政命令编制了人工智能应用案例清单，并提供了非机密和非敏感的当前及计划中的人工智能应用案例的公开清单。该清单详细说明了 HHS 如何利用人工智能，并包括监督方法和预期收益。2024 年，人工智能应用案例清单涵盖了 13 个机构的 271 个应用案例。第 13960 号行政命令启动了这个应用案例库，随后第 14110 号行政命令对其予以认可并进一步完善。HHS 将按照 OMB 备忘录 M - 24

- 10 的新要求，每年更新该清单。

HHS 近期优先事项：

除了对部门的人工智能应用案例进行编目外，HHS 还打算建立必要的内部流程和支持架构，以推动规范的人工智能应用。为此，HHS 的 OCAIO 将采取以下行动：

- 协调制定全部门的人工智能采购方法和工具包：这项工作将在 HHS 层面提供指导方针，以促进使用标准化的方法采购人工智能工具、技术和专业知识。该工作将与 HHS 各部门密切合作设计，以确保提供适用于整个 HHS 的充分指导，并与《联邦信息技术采购改革法案》的要求保持一致。HHS 的 OCAIO 还将探索在 HHS 采购法规和其他相关政策中纳入人工智能相关条款。
- 支持规范的原型制作和试点项目：这包括在部门和部门层面建立、共同领导并资助试点项目，以解决适用于多个 HHS 部门的通用解决方案以及特定使命的独特应用。HHS 的 OCAIO 旨在协助建立和使用“人工智能沙盒”，用于在跨部门部署前进行快速原型制作和解决方案评估（例如测试算法是否能产生预期结果）以及安全评估（例如部署算法是否会加剧或给 HHS 带来新的网络安全风险）。
- 确保对人工智能质量监测的监督：最终，HHS 及其各部门将负责确保人工智能符合适用标准。HHS 的 OCAIO 将在部门层面实施人工智能解决方案监测系统（包括准确性、可靠性和可追溯性），并推进和支持监测人工智能工具的能力建设。OCAIO 还将根据需要向 HHS 各部门发布指南，以帮助各部门建立适用于其机构使用的监测系统。
- 确保监督并更新流程以提升人工智能安全性：与上述质量监测不同，HHS 的 OCAIO 将与首席信息官办公室 (OCIO) 合作，确保人工智能应用案例符合适用的安全要求。OCIO 将按照其职责和流程，确保使用人工智能的信息技术得到妥善保护，并根据人工智能领域的变化需要更新安全流程。OCAIO 和 OCIO 将与 HHS 的其他利益相关者合作，确保这些流程能够在 HHS 各部门中得到应用。
- 发布人工智能使用指南：HHS 的 OCAIO 将提供指南，帮助 HHS 各部门确定在何种情况下使用人工智能解决方案最为合适。这些指南还将包括针对影响权利和安全的人工智能应用案例以及其他必要的人工智能应用案例所需采取的具体步骤，以符合行政命令、OMB 的 M - 24 - 10 备忘录及其他相关指南的要求。

7.5 劳动力与人才管理

背景：

打造一支具备人工智能能力的劳动力队伍，旨在让员工能够安全有效地履行职责，并在合理的情况下借助人工智能工具辅助工作。HHS 将持续评估在日常工作流程中利用人工智能的机会，并致力于对部门内动态变化的技术环境做出及时响应。在某些情况下，人工智能能够让员工将时间重新分配到最重要的领域，例如减少花在手动数据分析上的时间，而将更多时间用于决策制定和项目改进。

现有联邦行动：

其他联邦机构已经将利用人工智能赋能劳动力和人才管理列为优先事项，主要包括：

- 2024 年 2 月，人事管理办公室（OPM）制定了“未来劳动力”行动手册，详细阐述了劳动力战略，并为联邦机构提供了关于整合人工智能的指导方针。特别是，该行动手册为联邦机构提出了多项行动呼吁，包括将适当的人工智能能力融入人力资源流程、了解人工智能对劳动力的影响、提升团队的相关能力，以及对劳动力进行人工智能应用案例培训。
- 2023 年 12 月，OPM 在直接雇佣授权（DHA）框架中纳入了人工智能相关岗位，使联邦机构能够绕过特定的招聘流程，招聘高需求领域的人才。这一战略决策使各机构能够吸引和聘用具备专业技能的人工智能专家，满足机构的复杂需求，同时避免了可能阻碍他们加入政府机构的传统招聘程序。在此授权基础上，HHS 制定了标准化的人工智能职位描述，以加快通过 DHA 在全部门的招聘速度。
- 进行生成式人工智能试点项目，使劳动力能够实现此前手动进行的数据分析自动化，从而为决策制定和项目改进提供支持。ASTP 的首席人工智能官办公室和政策办公室正在探索生成式人工智能如何简化联邦规则制定过程中管理、分析和纳入公众意见的端到端流程。重点在于使用经过设计的提示来为政策办公室的规则制定活动生成可用的意见摘要。

HHS 近期优先事项：

为打造一支具备人工智能能力的劳动力队伍，HHS 的 OCAIO 将采取以下行动：

- 与政府各部门领导合作制定人工智能招聘战略：HHS 的 OCAIO 将与其他政府利益相关者（包括 OPM 和管理和预算办公室）合作，制定招聘具备专业技能的人工智能专家的战略。该战略可能包括确定部门层面的人工智能需求、评估人工智能相关岗位的薪酬标准，以及建立可供联邦各实体共享的资源（例如人工智能相关职位描述）。
- 与 HHS 内部领导合作进行人工智能技能差距评估：HHS 的 OCAIO 将与 HHS 的首席人力资本官和其他部门的劳动力领导合作，对部门当前人员的人工智能能力差距进行评估。这将确定需要有针对性干预的领域，可能包括提升现有人才的技能，或招聘具备这些技能的新人才（无论是在联邦政府内部还是从外部招聘）。差距评估的结果还将为缩小已识别差距的资金计划提供依据。
- 提高所有 HHS 员工的人工智能素养：除了进行差距分析外，HHS 的 OCAIO 将推动开展基础人工智能素养培训，帮助所有 HHS 员工和承包商更好地理解人工智能，了解人工智能技术的潜在益处、局限性和风险。

7.6 结论

在本章中，HHS 概述了该部门已经采取以及未来将继续采取的措施，以实现人工智能在

其内部运营中的益处，并在快速发展的人工智能领域保持灵活和与时俱进。HHS 认识到人工智能的变革潜力不仅体现在其各部门的工作以及健康和公众服务生态系统的众多利益相关者的工作中，也延伸到了部门自身的内部运营。HHS 认为人工智能在改善其面向公众的项目和服务、改进支持部门创新的流程、为政策和指南提供参考以及提高劳动力效率等方面具有巨大机遇。如果能够规范地把握这些机遇，将使 HHS 能够更好地履行其改善美国人民健康和福祉的使命。

结论

美国卫生与公众服务部（HHS）立志成为创新和使用规范人工智能的全球领军者，为全体美国人的健康福祉带来前所未有的提升。本战略规划阐述了 HHS 为达成这一目标拟采取的行动方案。

在快速的技术进步推动下，人工智能在医学研究与发现、医疗产品开发、安全性与有效性评估、医疗服务提供、公众服务供给、公共卫生、网络安全以及 HHS 自身运营等领域的应用，已从设想变为现实。近些年来，人工智能融入日常生活，在健康、公众服务和公共卫生生态系统中占据重要地位。这一转变体现为，人工智能可作为工具，助力提供高质量医疗服务、简化药物研发流程、加快并优化健康及公众服务领域的沟通交流。此外，人工智能还能推动健康公平，比如为有语言障碍的人群提供实时自动翻译服务，或是通过优化语音模式、流畅对话，为残障人士提供帮助。正如本战略规划各章节所探讨的，人工智能的应用带来诸多充满前景的益处，然而也伴随着一系列风险，例如人工智能可能传播偏见、对患者需求分类错误，或泄露机密信息。

HHS 不仅致力于推动人工智能的应用以实现更好的成果，还致力于保护患者、护理人员及所有利益相关者，使其免受本战略规划各章节所讨论的潜在风险的影响。这一承诺意味着要实施强有力的措施，应对这些挑战，同时挖掘人工智能的变革潜力。

随着人工智能的持续快速发展，HHS 将采取同样动态的应对方式，不断迭代本规划以及整体的人工智能相关工作，以便走在发展前沿，解决新出现的问题。这种积极主动的立场将涉及持续的效益与风险评估、利益相关者参与，以及实施严格的保障措施，确保人工智能的使用符合伦理道德且公平公正。HHS 还将持续挖掘各领域内及跨领域的大胆机遇与合作项目，力求改善人们的生活。HHS 的各个部门将继续发挥关键作用，发布指导方针和政策、分配资源、开展宣传和教育项目，并培养专业人才队伍。

HHS 鼓励社区合作伙伴、州和地方政府（STLT）以及其他公共和私营部门的合作伙伴，以规范的态度率先开发和应用人工智能，为美国人改善健康和公众服务。HHS 致力于与利益相关者合作，在本战略规划所详述的行动基础上继续推进工作，解决健康、公众服务和公共卫生领域面临的问题，同时通过所述的保障措施确保人工智能的安全、合理使用。HHS 将继续支持与合作伙伴的互动和信息透明化，推动创造具有实际影响力的以人为本的解决方案。

HHS 旨在继续引领健康、公众服务和公共卫生领域的创新，满足美国民众不断变化的需求，本规划只是支撑该部门应对未来挑战的基础性一步。HHS 矢志支持能够增进全体美国人健康福祉的人工智能技术。

附录 A：术语词汇表

表格 1：关键术语词汇表

术语	定义
人工智能问责制	人工智能系统的创造者应就人工智能系统的结果负责，包括对造成的任何损害进行补偿的原则
人工智能伦理	研究人工智能的道德影响和社会影响的伦理学分支
人工智能赋能医疗器械、人工智能赋能设备和 / 或人工智能设备	在本规划中，“人工智能 / 机器学习赋能医疗器械”“人工智能赋能设备”和“人工智能设备”这几个术语可互换使用，指的是：(1) 能够执行医疗器械功能（例如诊断、治愈、缓解、治疗、预防），且不属于传统硬件医疗器械一部分的人工智能软件；(2) 作为医疗器械一部分或不可或缺组成部分的人工智能软件
人工智能	根据第 14110 号行政命令第 3 (b) 条以及《美国法典》第 15 编第 9401 (3) 条，人工智能是一种基于机器的系统，能够针对一组给定的人类定义目标，进行预测、推荐或决策，从而影响真实或虚拟环境。人工智能系统利用基于机器和人类的输入来感知真实和虚拟环境；通过自动化分析将这些感知抽象为模型；并利用模型推理制定信息或行动选项
人工智能性能监测	指定期收集和分析已部署人工智能系统的使用数据，以评估其在现实环境中完成预定任务的性能的过程。对人工智能模型性能的评估涉及各种性能指标和标准，具体取决于特定应用。这种监测通常旨在评估这些人工智能系统在实际应用中的性能，检测性能下降或变化（例如由于数据漂移），识别滥用情况，并解决任何安全或可用性问题
人工智能系统	任何全部或部分使用人工智能运行的数据系统、软件、硬件、应用程序、工具或实用程序
辅助型人工智能	旨在辅助人类决策的人工智能产品。人工智能仅提供建议、信息或数据，帮助用户做出更明智的决策。辅助型人工智能和自主型人工智能处于一个逐渐过渡的范围内。辅助型人工智能的示例可能包括监测患者生命体征并在某些指标超出正常范围时向用户或医疗保健提供者发出警报的可穿戴设备，或通过显示潜在畸形来辅助放射科医生的产品

术语	定义
自主型人工智能	能够独立执行任务、运作并做出决策而无需人类干预的人工智能产品，例如人工智能代理。其自主程度因产品而异。辅助型人工智能和自主型人工智能处于一个逐渐过渡的范围内。自主型人工智能的一个例子可以是一种能够自动识别正常 X 光片并生成报告而无需放射科医生介入的产品
人工智能中的偏差	在人工智能算法和数据集中引入的偏见性假设和偏好，这可能导致不公平的结果或决策
生物制品	根据《公共卫生服务法》，“生物制品”一词指病毒、治疗血清、毒素、抗毒素、疫苗、血液、血液成分或衍生物、变应原产品、蛋白质或类似产品，或肺炎链球菌或肺炎链球菌衍生物（或任何其他三价有机砷化合物），适用于人类疾病或病症的预防、治疗或治愈
聊天机器人	一种程序，通过文本或语音命令实现大语言模型与人类之间的沟通，模拟人与人之间的对话
临床决策支持软件(CDS)	旨在为疾病或其他病症的诊断、治疗、预防、支持或治愈提供决策支持的软件
云计算	一种模式，旨在实现随时随地、方便快捷的按需网络访问，能够连接到由可配置的计算资源（例如网络、服务器、存储、应用程序和服务）组成的共享资源池。这些资源可以快速配置和释放，且只需最少的管理操作或服务提供商的沟通
人工智能虚构	人工智能模型在接收到准确数据的情况下生成虚假或误导性信息的现象
持续机器学习	模型能够通过随时间纳入新数据或经验来调整其性能，同时保留先前的知识 / 信息。模型的变化通过一个定义明确的过程实施和验证，该过程旨在基于对新数据的分析来提高性能。与锁定模型不同，持续机器学习模型具有定义明确的学习过程来改变其行为

术语	定义
卷积神经网络 (CNN)	一种专门的深度神经网络架构，由一个或多个卷积层组成，适用于处理网格状数据，如图像。在卷积层中，“滤波器”（窗口或模板）在输入图像的区域上滑动，通过应用卷积（一种对输入数据进行的数学点运算）来识别低级模式（例如边缘）。可以应用不同的滤波器来提取不同的特征，如图像中的边缘、纹理或曲线。此外，卷积神经网络可以包括池化层，其功能是在保留相关特征的同时降低特征维度。这些卷积层和池化层相互堆叠，使网络能够对模型建立层次化的理解，这使得卷积神经网络在图像识别和计算机视觉等任务中非常有效。该网络的一个重要特点是它能够在进行特征提取的同时保留原始输入的空间信息
数据卡	一份关于数据集相关特征的结构化报告，旨在满足利益相关方对人工智能开发和评估的需求。它包含一个描述性部分，包括样本数量、收集协议及相关元数据等描述性信息，以及一个评分卡部分，该部分使用相关标准和指标对数据集特征进行定量分析报告
数据漂移	指已部署模型接收到的输入数据分布随时间发生变化，这可能导致模型性能下降。当基础数据的属性发生变化时就会出现数据漂移。数据漂移会影响预测模型的准确性和可靠性。例如，医疗人工智能产品可能会由于模型开发所用数据与临床操作所用数据之间的统计差异（由于医疗实践的差异或训练与临床使用之间的使用情境不同），以及患者人口统计学、疾病趋势和数据收集方法随时间的变化而经历数据漂移
数据治理	基于内部数据标准和政策，管理企业系统中数据的可用性、易用性、完整性和安全性的过程，这些标准和政策也控制数据的使用
数据隐私	信息技术领域中涉及组织或个人确定计算机系统哪些数据可以与第三方共享的方面
数据标准	一种标准类型，是一种商定的方法，用于实现对对象、过程或信息单元的一致测量、鉴定或交换。数据标准指的是组织、记录和格式化数据的方法，以辅助数据的聚合、共享和重用

术语	定义
数据使用协议 (DUA)	拥有数据源（通常是数据集或数据库）访问权限的实体与将接收数据或其子集以供再次使用的次要实体之间的法律合同。数据使用协议概述了共享数据的使用条款和限制，次要实体可能需要满足某些标准，例如其所属机构、教职员身份以及其研究项目的机构审查委员会（IRB）批准。限制示例包括限制对共享数据的访问、要求任何研究成果的传播都要引用数据及其来源实体、要求在研究期结束时销毁数据文件，以及对数据用于商业目的的限制。访问包含受保护健康信息（PHI）的数据通常需要数据使用协议
数据驱动的人工智能	强调数据在增强技术从人类智能中学习并进行增强的能力方面重要性的人工智能。它涉及有效的数据理解、治理，以及一种通过人工智能扩展数据价值以增强业务流程的思维方式
深度学习	机器学习的一个专门分支，涉及训练具有多个中间（隐藏）层的神经网络，这些层在接收数据的输入层和呈现最终网络输出的输出层之间运行。每一层都学习将其输入数据转换为更抽象和更复合的表现形式，并产生作为下一层输入的输出。随着数据在连续的层中传播，这些模型可以从输入数据中学习层次化的特征表示。例如，在医疗保健领域，深度学习模型可用于识别医学图像中的肿瘤或可疑病变，以支持医生和放射科医生进行疾病评估
深度伪造	看似真实但已通过人工智能技术进行操纵的视频、照片或音频记录。底层技术可以替换人脸、操纵面部表情、合成人脸和合成语音。深度伪造可以描绘某人似乎在说或做他们实际上从未说过或做过的事情
数字健康技术 (DHT)	一种使用计算平台、连接性、软件和 / 或传感器用于医疗保健及相关用途的系统。这些技术的应用范围广泛，从一般健康应用到作为医疗器械的应用。它们包括旨在用作医疗产品、用于医疗产品或作为其他医疗产品（设备、药物和生物制品）辅助工具的技术。它们也可用于开发或研究医疗产品

术语	定义
数字孪生	一组信息结构，能够模拟物理性质的结构、情景和行为，在其整个生命周期中通过来自其物理孪生的数据进行动态更新，并为决策提供信息。虚拟与物理之间的双向交互是数字孪生的核心。数字孪生可实现个性化医疗应用。例如，患者的数字孪生可以为临床决策提供信息，例如治疗方案和临床评估。此外，数字孪生可以在组建大型、多样化的虚拟人群队列以进行计算机模拟临床试验，以及在药品生产过程的质量评估和流程优化中发挥作用
药物	根据《联邦食品、药品和化妆品法案》，“药物”一词指：(A) 在《美国药典》《美国顺势疗法药典》或《美国国家处方集》或其任何补充剂中认可的物品；(B) 旨在用于人类或其他动物疾病的诊断、治愈、缓解、治疗或预防的物品；(C) 旨在影响人类或其他动物身体结构或任何功能的物品（食品除外）；(D) 旨在用作 (A)、(B) 或 (C) 款中指定物品的成分的物品
集成方法	一种机器学习技术，结合多个模型以提高整体预测性能，相比使用单个模型更具优势。这涉及训练一组基础模型，例如神经网络，然后聚合它们的预测以做出最终预测。一些常见的集成方法包括引导聚合算法（即使用训练数据的不同子集训练多个模型并平均它们的预测）、提升法（即按顺序训练模型，每个新模型专注于纠正前一个模型的错误）和堆叠法（即使用多个基础模型的预测作为更高层次“元模型”的输入特征，该元模型能够学习如何最好地组合它们）
可解释性	“指对人工智能系统运行背后机制的描述”（来源：美国国家标准与技术研究院）。可解释性可能有助于克服黑箱系统（即内部运作和决策过程不透明或不易理解的系统）的不透明性。这些解释可以使用多种形式，包括自由文本解释、显著图、Shaply 加性解释（SHAP）或来自数据的相关输入示例。其主要目的是回答“为什么”人工智能系统做出特定决策。适当的可解释人工智能（XAI）方法可以帮助开发更准确、公平、可解释和透明的人工智能系统，以安全地辅助医疗保健中的人类决策
探索性数据分析(EDA)	一种在做出进一步假设或测试假设之前分析数据集以总结其主要特征的方法，通常使用可视化手段

术语	定义
特征工程	一种从原始数据中识别出最能代表潜在模式的属性，以用于训练特定机器学习模型的机器学习过程。它涉及选择、转换或创建相关的输入变量（称为特征），以提高机器学习模型的性能。可以利用领域知识和数据分析技术来构建能够捕捉数据中内在关系的特征。例如，对于一个可以预测心力衰竭的模型，对患者数据进行特征工程可能涉及通过组合相关特征（如年龄、血压、胆固醇水平和心血管疾病史）创建一个“风险评分”
联邦学习	一种分散式的训练机器学习模型的方法。模型由每个站点在本地保存的数据上进行训练，模型更新被发送到中央服务器，中央服务器聚合这些更新以改进全局模型。这种方法旨在保护数据隐私，因为原始数据保留在本地站点，不会被集中。例如，联邦学习可以让医院在不共享患者数据的情况下协作开发心脏病预测模型。模型被发送到每个医院进行本地训练，只有每个医院的模型更新（而不是原始数据）被发送回并进行聚合。这样，个人患者信息保持本地化，在解决隐私问题的同时仍能受益于共同改进的模型
基础模型	使用大型、通常未标记的数据集和大量计算资源进行训练的人工智能模型，适用于广泛的场景，包括一些模型未专门开发和训练的场景（即具有涌现能力）。这些模型可以作为基础，通过进一步训练（即微调）为特定用途构建和调整其他模型。这些模型可以执行一系列通用任务，如文本合成、图像操作和音频生成。这些模型基于深度学习架构，如 Transformer，可以使用单模态或多模态输入数据
生成对抗网络 (GAN)	一种基于深度学习的模型架构，通常由两个相互竞争的神经网络组成，即生成器和判别器。生成器的目标是合成假数据以欺骗判别器，而判别器试图区分合成示例（生成器的输出）和原始训练数据分布。训练的目标是在两个竞争网络之间找到平衡点，训练过程结束后，生成器学会生成与训练集分布相同的新数据。这种方法可用于生成合成图像
生成式人工智能 (GenAI)	“一类模拟输入数据的结构和特征以生成衍生合成内容的人工智能模型。这可以包括图像、视频、音频、文本和其他数字内容”（来源：第 14110 号行政命令）。这通常通过近似输入数据的统计分布来实现。例如，在医疗保健领域，生成式人工智能可用于在合成医疗数据上生成注释（例如图像特征、文本标签），以帮助扩展用于训练算法的数据集

术语	定义
健康信息交换 (HIE)	健康信息交换允许医疗保健专业人员和患者适当地访问并安全地以电子方式共享患者的医疗信息。如今，不同形式的健康信息交换背后的技术受到许多医疗服务提供场景的驱动
人机回环机器学习	一种人类与机器学习模型进行交互以提高准确性和终端用户对机器信任的方法。在人机回环机器学习中，人类交互是迭代的，并且可以随着时间的推移导致性能的持续改进。这种交互在模型对其预测不确定并需要人类指导进行验证的场景中尤为重要。与有监督的机器学习不同，有监督的机器学习主要在数据标记阶段涉及人类输入，之后算法独立训练。标记或注释是将描述性信息附加到数据的过程。在注释过程中，数据本身不会改变
以人为本的人工智能(HCAI)	强调人工智能技术对个人和社会的影响，优先考虑人类福祉、需求和目标的人工智能
互操作性	在不同的信息技术系统、软件应用程序和网络中，能够准确、有效、安全和一致地进行通信和交换数据的能力，并且数据交换能够保留数据的临床或操作目的及含义，不发生改变
关键绩效指标 (KPI)	一种可衡量的值，用于展示组织在实现关键业务目标方面的有效程度
大语言模型	一种人工智能模型，在大型文本数据集上进行训练，以学习自然语言中单词之间的关系。这些模型可以应用所学模式，对收到的广泛输入或提示预测生成自然语言响应，以执行诸如翻译、总结和问答等任务。这些模型的特点是具有大量的模型参数（即训练模型中的内部学习变量）。大语言模型基于基础人工智能模型进行构建，通过开发超越基本语言模式获得更全面的语言理解能力。例如，在大语言模型的背景下，聊天机器人是一种通过文本或语音命令实现大语言模型与人类之间通信的程序，模拟人与人之间的对话
锁定模型	每次应用相同输入时都提供相同输出，并且其参数或配置不能更新，不会随使用而改变的模型。对于人工智能赋能的医疗产品，锁定模型有助于确保性能的一致性
机器学习(ML)	一组可用于训练人工智能算法以基于数据提高任务性能的技术

术语	定义
机器学习算法	执行任务或解决问题所遵循的逐步程序或一组指令。例如，在机器学习中，算法用于使用数据训练模型以解决特定问题
机器学习算法偏差	“偏差”一词在不同领域和行业的各种背景下使用。在人工智能的背景下，偏差指的是模型预测或结果针对某些数据点或群体与其他数据点或群体相比存在的系统性偏差。这里我们关注的是算法偏差，这种偏差可能源于多种来源，例如训练数据集的特征、模型开发过程中的选择、数据处理的不规则性，或数据收集过程中引入的偏差或人类决策。算法偏差可能导致对某些对象、人或群体与其他对象、人或群体相比在处理上存在系统性差异或错误，或导致预测失败，进而可能产生其他风险，其中处理指的是任何行动，包括感知、观察、表示、预测或决策
机器学习模型	一种数学构造，能够对输入数据生成推断或预测。该模型是机器学习算法从数据中学习的结果。模型由算法训练，算法是用于处理数据并得出结果的逐步程序。人工智能系统（例如人工智能赋能的医疗器械）使用一个或多个模型来实现其预期目的
医疗器械	根据《联邦食品、药品和化妆品法案》，“设备”指仪器、器械、工具、机器、装置、植入物、体外试剂或其他类似或相关物品，包括任何组件、部件或配件，其：(A) 在《美国国家处方集》或《美国药典》或其任何补充剂中被认可；(B) 旨在用于人类或其他动物疾病的诊断、治愈、缓解、治疗或预防；(C) 旨在影响人类或其他动物身体的结构或任何功能，并且其主要预期目的不是通过在人体或其他动物体内或体表的化学作用实现，且不依赖于被代谢来实现其主要预期目的。“设备”一词不包括第 520 (o) 条规定的软件功能
医疗产品	在本规划中，“医疗产品”一词统称本词汇表中定义的药物、生物制品和医疗器械（包括一些基于软件的行为干预措施）
指标	用于跟踪和评估特定过程、项目或活动状态的定量测量
模型校准	调整机器学习模型生成的预测概率，以确保它们准确反映现实世界中事件或结果的观察频率的过程。例如，如果一个模型校准良好，并且预测某患者患乳腺癌的概率为 20%，那么在该模型做出此类预测的 100 名患者中，实际患乳腺癌的患者数量应大致为 20 人

术语	定义
模型卡片	一份关于人工智能模型相关技术特征以及在各种条件下（例如不同文化、人口统计或表型群体以及与预期应用领域相关的交叉群体）的基准评估结果的结构化报告。模型卡片还提供有关模型预期使用环境的信息以及评估其性能的详细信息
模型部署	将机器学习模型集成到现有的生产环境中，以进行实际可行的预测的过程
模型拟合	训练机器学习模型以通过调整训练参数捕捉数据中的潜在模式，使模型的预测尽可能接近训练数据中的目标值的过程。参数的调整使模型能够对数据形成一般性的理解，使其在对新的、未见过的数据进行预测时有用。一个拟合良好的模型既不会过拟合也不会欠拟合，而是在训练数据和新的、未见过的数据上都表现良好，因为它正确地捕捉了输入和目标变量之间的关系
模型稳健性	机器学习模型在不同情况下保持其目标或指定性能水平的能力。这些情况可能包括噪声数据（例如包含错误、不一致和缺失值的数据）、未见过的数据或数据漂移，或旨在欺骗模型的对抗性攻击。例如，在医疗保健领域，模型稳健性的挑战可能出现在医学图像分类中，成像条件的变化（如光照或分辨率）可能会影响基于标准化图像训练的肿瘤分类模型的性能
模型权重	人工智能模型中的一个数值参数，有助于确定模型针对输入的输出
多模态	一种处理和整合多种不同数据类型的方法，旨在捕捉和利用它们之间的关系，以更好地理解输入信息或提高预测性能。这些数据类型可能包括文本、图像、音频、视频、基因组学、传感器数据等。这些不同的数据类型可以使用单个多模态网络（例如基于神经网络或其他架构）进行处理，也可以通过单独的单模态网络（例如用于文本的大语言模型和用于图像的卷积神经网络）进行处理，其中单模态输出被组合在一起。例如，在医疗保健领域，电子健康记录和可穿戴生物传感器的数据可以结合起来，实现对患者的远程监测

术语	定义
国家生命统计系统(NVSS)	国家生命统计系统是公共卫生领域政府间数据共享历史最悠久且最成功的范例，其共享关系、标准和程序构成了国家卫生统计中心 (NCHS) 收集和传播美国官方生命统计数据机制。这些数据通过 NCHS 与负责生命事件（出生、死亡、婚姻、离婚和死产）登记的各个司法管辖区运营的生命登记系统之间的合同提供
自然语言处理 (NLP)	人工智能和语言学的一个子领域，使计算机能够理解、处理、解释和生成人类语言。自然语言处理系统可以使用计算语言学和机器学习技术处理和分析自然语言数据，执行文本分类、情感分析和翻译等任务。自然语言生成是自然语言处理的一个应用，涉及使用人工智能系统生成人类可读的文本输出，如摘要、报告、故事或回复
神经网络	一种受人类大脑结构启发的计算模型。它由相互连接的节点（或“神经元”）组成，这些节点被组织成层：接收数据的输入层、一个或多个处理并识别数据中模式的隐藏层，以及呈现最终网络输出的输出层
过拟合	在机器学习中，过拟合发生在模型过度学习训练数据，不仅捕捉到基本模式，还捕捉到噪声或随机波动时。这样的模型在训练数据上可能表现出色，但在对新的、未见过的数据进行泛化时会遇到困难
性能指标	在人工智能领域，性能指标是用于评估模型在给定任务中产生期望输出能力的定量或定性测量。指标的选择取决于具体任务和模型目标。定量指标的示例包括准确率、精确率、敏感度（召回率）、特异度、F1 分数和受试者工作特征曲线下面积 (AUC - ROC)。定性测量可能涉及热图评估或可视化解释。这些指标能够对模型进行系统评估、比较和优化，并有助于评估模型是否达到其预期目标
个人身份信息 (PII)	机构保存的关于个人的任何信息，包括：(1) 任何可用于区分或追踪个人身份的信息，如姓名、社会安全号码、出生日期和地点、母亲的婚前姓氏或生物识别记录；(2) 任何其他与个人相关或可关联到个人的信息，如医疗、教育、财务和就业信息

术语	定义
药物警戒	根据 FDA 的《行业指南：良好药物警戒实践和药物流行病学评估》，该指南适用于与药物和生物制品（不包括血液和血液成分）相关的活动，“药物警戒”一词指“与检测、评估和理解不良事件相关的所有科学和数据收集活动。这包括使用药物流行病学研究。这些活动的目的是识别不良事件，并在可能的范围内了解其性质、频率和潜在风险因素”
预测分析	使用数据、统计算法和机器学习技术，基于历史数据识别未来结果可能性的方法
人工智能中的隐私	在人工智能系统的开发和应用中，对个人数据和信息的保护，确保数据的使用符合伦理道德且获得同意
隐私增强技术	任何用于减轻数据处理过程中产生的隐私风险的软件或硬件解决方案、技术流程、技术或其他技术手段，包括通过增强可预测性、可管理性、可分离性、存储、安全性和保密性来实现。这些技术手段可能包括安全多方计算、同态加密、零知识证明、联邦学习、安全飞地、差分隐私和合成数据生成工具
概念验证(PoC)	项目开发的早期阶段，展示一个想法或技术在解决特定问题方面的可行性，以证明其潜在应用
受保护健康信息(PHI)	由受保实体或其业务合作伙伴以任何形式或媒介传输或维护的个人可识别健康信息（《联邦法规汇编》第 45 编第 160.103 条）。该定义排除了少数几类个人可识别健康信息，例如受保实体作为雇主持有的就业记录中包含的个人可识别健康信息
阅读理解与生成(RAG)	一种人工智能技术，通过提供一个数据库作为参考，增强对文本的理解和生成能力，旨在避免语言模型出现幻觉等问题
（人工智能中的）参考标准	确定或测量所研究现象真实状态或属性的最佳可用方法，在人工智能中通常以标记数据的形式表示。它作为一个基准，用于评估模型的输出。在临床环境和医学研究中，参考标准是一种临床诊断的金标准测量或方法，用于验证结果。例如，参考标准可以指示疾病或异常的存在、程度和位置。标记或注释是将描述性信息附加到数据的过程。在注释过程中，数据本身不会改变

术语	定义
强化学习	一种机器学习方法，其中模型（或智能体）通过与环境的交互采取行动并获得奖励或惩罚来进行学习。模型从其行动的后果中学习，而不是通过明确的教导，并且根据过去的经验（利用）和做出新的选择（探索）来选择行动，本质上是一种试错学习。例如，在医疗保健领域，强化学习可用于为慢性病患者推荐个性化治疗方案。模型接收患者数据，包括他们的病史、当前健康状况和治疗反应，然后提出治疗方案。关键在于反馈循环：随着患者数据不断更新其对治疗的反应信息，模型会相应地调整其建议。这个过程涉及大量的试错，因为模型从每次患者交互中学习。随着时间的推移，通过许多这样的交互，模型变得更擅长为个体患者预测和推荐最有效的治疗方案
人工智能可靠性	人工智能系统在特定条件下持续运行，提供准确可靠结果的能力
规范的人工智能(RAI)	秉持社会道德价值观的人工智能实践，确保人工智能系统公平运行、按预期发挥作用，并对其结果负责。这包括遵守公平、透明、问责、安全、隐私和可靠等原则
人工智能稳健性	人工智能系统在面对变化的条件或处理意外或对抗性输入时保持其性能的能力
沙盒	一个安全、受控、受限的环境，允许在不受特定法规约束的情况下测试产品、监管方法和其他技术，这些法规在安全、受控、受限的沙盒环境之外通常是不允许的
可扩展且可互操作的人工智能	确保在组织内的使用高效、适应性强，并与现有工作流程相协调的人工智能，使基于人工智能的解决方案能够与机构目标同步发展和运行
自监督机器学习	机器学习算法能够从可用的未标记数据中生成自己的标签。监督学习不同监督学习需要提供标记数据，而无监督学习在没有标签的情况下发现隐藏模式，自监督学习利用数据中的固有结构创建自己的标签。当标记数据有限或不可用时，这种方法很有用

术语	定义
半监督机器学习	结合了无监督和有监督技术的机器学习算法。有监督学习技术使用标记数据进行训练，而无监督学习技术使用未标记数据进行训练。标记或注释是将描述性信息附加到数据的过程。在注释过程中，数据本身不会改变。例如，考虑从胸部 X 光片中诊断肺部疾病的任务。半监督学习模型首先在一小部分已标记的 X 光图像上进行训练，这些图像已被放射科医生标记为显示特定肺部疾病迹象或正常。然后，模型利用这些知识开始对更大的未标记图像集进行预测
监督机器学习	使用标记数据进行训练的机器学习算法。标记或注释是将描述性信息附加到数据的过程。在注释过程中，数据本身不会改变
合成数据	通过人工创建（例如通过统计建模、计算机模拟）的数据，从而生成新的值和 / 或数据元素。一般来说，合成数据旨在表示实际患者数据中看到的结构、属性和关系，但不包含关于个人的任何真实或特定信息。例如，在医疗保健领域，合成数据是人工生成的数据，旨在模仿真实患者数据的属性和关系。合成数据是部分或完全使用计算技术生成的示例，而不是通过物理系统从人类受试者获取的
测试数据	这些数据用于表示人工智能系统的性能。在训练过程中，算法从未见过这些数据，它们用于在训练后估计人工智能模型的性能。在系统部署或上市之前进行测试，以生成证据来确定人工智能系统的性能。对于人工智能赋能的医疗产品，测试数据应独立于用于训练和调整的数据
测试平台	一个配备了用于对工具和技术（包括人工智能和隐私增强技术）进行严格、透明和可重复测试的设施或机制，以帮助评估这些工具或技术的功能、可用性和性能
训练数据	人工智能系统的制造商在构建人工智能模型的程序和训练算法中使用的数据，包括用于定义模型权重、连接和组件的数据

术语	定义
迁移学习	机器学习中的一种策略性方法，其中为特定任务开发的模型被适配用于第二个任务。这种方法利用从先前解决的问题（源任务）中获得的知识和模式，来提高模型在后续（通常是类似的）问题（目标任务）上的性能和学习效率。例如，在医疗保健领域，一个经过训练用于识别肺部 X 光图像中肿瘤的模型，可以利用学到的模式来提高对肝脏超声图像中异常的识别能力
人工智能的透明度和可解释性	人工智能系统能够被人类理解，并且其过程和结果能够以人类可理解的方式进行解释的能力
调整数据	这些数据通常由人工智能系统的制造商用于评估少量经过训练的模型。这个过程涉及探索各个方面，包括不同的架构或超参数（即用于针对任务调整模型的参数）。调整阶段发生在人工智能系统测试阶段之前，是训练过程的一部分。虽然人工智能和机器学习社区有时使用“验证”一词来指代调整数据和调整阶段，但 FDA 通常不会在这种情况下使用“验证”一词，因为其在监管上有特定定义（见《联邦法规汇编》第 21 编第 820.3 (z) 条）
欠拟合	在机器学习中，欠拟合发生在模型未能捕捉到训练数据的模式和复杂性时，导致在训练数据和新的、未见过的数据上的性能都较差
无监督机器学习	在训练过程中仅使用未标记数据的机器学习算法。无监督学习旨在发现数据中隐藏的模式或结构
用户体验设计	以关注用户与产品交互的质量和效率为核心，设计产品、系统或服务的过程
用户研究	通过观察技术、任务分析和其他反馈方法，了解用户行为、需求和动机的研究
水印嵌入	将信息嵌入人工智能生成的输出（包括照片、视频、音频片段或文本等输出）中的行为，这些信息通常难以去除，目的是验证输出的真实性，或其来源、修改或传播的身份或特征

表格 2：首字母缩写

术语	全称
ACF	儿童与家庭管理局
ACL	社区生活管理局
AHRQ	医疗保健研究与质量局
AI	人工智能
AIDR	人工智能数据准备度
ARPA - H	卫生高级研究计划局
ASPR	战略准备与响应管理局
ASTP/ONC	科技政策助理部长办公室 / 国家卫生信息技术协调员办公室
ATSDR	有毒物质与疾病登记署
CAIO	首席人工智能官
CBER	生物制品评估与研究中心
CDC	疾病控制与预防中心
CDER	药品评估与研究中心
CDRH	器械与放射卫生中心
CDS	临床决策支持
CGMP	现行药品生产质量管理规范
CMS	医疗保险与医疗补助服务中心
CPT	现行操作术语
CRDC	癌症研究数据共享库

术语	全称
DMI	数据现代化倡议
DOE	能源部
ECG	心电图
EHR	电子健康记录
FDA	食品药品监督管理局
FTC	联邦贸易委员会
GSA	总务管理局
HHS	卫生与公众服务部
HIPAA	健康保险流通与责任法案
HRSA	卫生资源与服务管理局
IDE	研究性器械豁免
IHS	印第安卫生服务局
IND	研究性新药
IRB	机构审查委员会
LEAP	前沿加速项目
LLM	大语言模型
ML	机器学习
MoA	作用机制
MRI	磁共振成像

术语	全称
NCHS	国家卫生统计中心
NIH	国立卫生研究院
NIST	美国国家标准与技术研究院
NLP	自然语言处理
NOFO	资金机会公告
NPSD	患者安全数据库网络
NTAP	新技术附加支付
OCAIO	首席人工智能官办公室
OMB	管理与预算办公室
PDSI	预测性决策支持干预措施
PHI	受保护的健康信息
PI	预测智能
PII	个人可识别信息
PoC	概念验证
PSO	患者安全组织
RCM	收入周期管理
SAMHSA	药物滥用和心理健康服务管理局)
SaMD	医疗器械软件
SDOH	健康的社会决定因素

术语	全称
TA	治疗领域
TPLC	产品全生命周期
TTS	技术转型服务
UDS	统一设计系统
XAI	可解释人工智能

附录 B：部分联邦政策和法规

表格 3：支持人工智能合理使用的非详尽联邦政策和法规

政策重点与目标	具体法规、政策或指南	简要描述
关于人工智能的总体立法和行政行动： 制定联邦在人工智能领域的协调方针，涵盖其在联邦政府内的影响，以改进美国的人工智能并确保其持续安全和规范的使用	第 14110 号行政命令（安全、可靠和值得信赖的人工智能开发与使用）	强调持续安全使用人工智能的重要性，并要求包括 HHS 在内的多个联邦机构制定人工智能战略
	《人工智能权利法案蓝图》	确定了五项原则，应指导自动化系统的设计、使用和部署，以在人工智能时代保护美国公众：安全有效的系统；算法歧视保护；数据隐私；通知与解释；人类替代方案、考量和后备措施
	《2020 年国家人工智能倡议法案》	呼吁在整个联邦政府开展统一项目，以加速人工智能研究与应用，促进国家经济繁荣和保障国家安全
	第 13859 号行政命令（保持美国在人工智能领域的领导地位）	定义了一个统一的联邦政府人工智能战略，聚焦于推动技术突破、制定适当的人工智能标准、培训当前和未来劳动力、培养公众信任和信心，以及营造支持美国人工智能研究的国际环境
	第 13960 号行政命令（促进联邦政府使用可靠的人工智能）	为联邦政府机构使用人工智能制定原则

政策重点与目标	具体法规、政策或指南	简要描述
	OMB M - 21 - 06 (人工智能监管应用指南)	为所有联邦机构提供指导，以制定关于人工智能相关技术和产业部门的监管和非监管方法，以及思考减少人工智能技术开发和使用障碍的方法。HHS 对此 OMB 指南做出回应，提供了授权 HHS 发布关于私营部门人工智能应用开发和使用法规的法定权限等内容
	第 1557 条 92.210 款 (患者护理决策支持工具中的非歧视规定)	保护个人在医疗项目或活动中，不因种族、肤色、民族、性别、年龄或残疾，在使用患者护理决策支持工具时受到歧视
研究参与者保护措施： 确立保护研究参与者福祉、隐私与自主权的预期目标和最优方法。这些政策、法规及最优方法中所包含的伦理考量（如隐私），能够解决人工智能在研究开发与应用过程中的关键问题。研究人员遵循这些准则，便可以减轻因人工智能使用与开发所产生的潜在危害和不公平现象。	《保护人类受试者》(《联邦法规法典》第 45 编第 46 部分)	为涉及人类参与者及其数据的 HHS 支持研究中的 IRB、知情同意和合规保证提供基本规定，包括对风险与收益的考量
	《保护人类受试者》(《联邦法规法典》第 21 编第 50 部分) 和《机构审查委员会》(《联邦法规法典》第 21 编第 56 部分)	规定了临床研究中合规和 IRB 的相关条款，这些临床研究同时也受 FDA 监管。\\
	保密证书	禁止将可识别的敏感研究信息披露给与研究无关的任何人，除非参与者同意或在少数其他特定情况下
	NIH 关于数据和生物样本二次研究知情同意书	提供了在研究计划中存储和共享收集的数据和生物样本以供未来使用时，知情同意文件的考虑要点、使用说明和可选示例语言
	《通用规则》	要求在将人类受试者纳入研究之前，获得具有法律效力的知情同意

政策重点与目标	具体法规、政策或指南	简要描述
	知情同意发布说明	提供关于如何遵守《通用规则》中在将人类受试者纳入研究前获得知情同意这一要求的一般指导
	NIH 关于在共享人类研究参与者数据时保护隐私的信息	在 NIH 支持的研究中共享数据时，提供保护人类研究参与者隐私的一系列原则和最优方法（根据 NIH 数据管理与共享政策发布）
患者保护： 助力保护健康数据的隐私与安全涵盖医疗服务、研究与发现等诸多领域。	《健康保险流通与责任法案》隐私规则	《健康保险流通与责任法案》有助于保护医疗研究中使用的健康数据的隐私和安全，包括涉及人工智能的研究，从而增强对医疗研究活动的信任。隐私规则规定了受保实体出于研究目的使用或披露受保护健康信息的条件
	健康数据、技术和互操作性：认证计划更新、算法透明性和信息共享 (HTI-1) 最终规则	实施《21 世纪治愈法案》的相关条款，并通过新的和更新的标准、实施规范和认证标准，对 ONC 健康信息技术认证计划（认证计划）进行更新。HTI - 1 最终规则中的条款促进了互操作性，提高了透明度，并支持电子健康信息的访问、交换和使用
	健康数据、技术和互操作性：可信交换框架与通用协议 (HTI-2) 最终规则	在 HTI - 1 最终规则基础上进行的技术和标准更新，范围从交换临床图像（例如 X 光）的能力到增加多因素身份验证支持
	《21 世纪治愈法案》	有助于加速医疗产品开发，使新的创新和进展能够更快、更高效地惠及有需要的患者。该法律基于 FDA 先前的工作，将患者的观点纳入 FDA 在药物、生物制品和器械开发决策过程中，并包含与隐私保护和确保适当获取电子健康信息相关的条款

政策重点与目标	具体法规、政策或指南	简要描述
	《健康信息技术经济和临床健康法案》	赋予 HHS 建立项目的权力，通过推广健康信息技术，包括电子健康记录和安全的私人电子健康信息交换，来提高医疗保健质量、安全性和效率。该法案解决了与电子健康信息交换相关的隐私和安全问题，包括更严格的数据泄露通知要求
生物安全与生物安保： 构建并作为全面生物安全与生物安保监督体系的一部分。由美国卫生与公众服务部（HHS）资助的研究，包括使用由人工智能赋能或辅助的工具与技术开展的研究，均在这一监督框架之内。尽管其中一些政策并未明确提及人工智能，但如果涉及生物制剂、毒素或核酸分子的研究包含受这些政策覆盖的实体实验，那么这些政策同样适用于此类研究中人工智能的开发与应用。	美国政府对两用性关注研究和具有增强传染病潜力病原体的监督政策（2025 年 5 月 6 日生效）	为开展和管理某些类型的联邦资助的生命科学研究，对可能对公共卫生、农业、粮食安全、经济安全或国家安全构成风险的生物制剂和毒素，提供一个统一的联邦监督框架。该政策“鼓励对可能导致开发出直接有助于设计‘具有增强传染病潜力的病原体或新型生物制剂或毒素’的计算模型的计算机模拟研究进行机构监督，无论资金来源如何”。一旦生效（2025 年 5 月 6 日），这个统一框架将取代当前通过以下方式划定的监督范围： ● 美国政府对生命科学两用性关注研究的监督政策； ● HHS 关于指导涉及具有增强传染病潜力病原体的拟议研究的资助决策框架
	美国政府核酸合成筛查框架（2024 年 10 月 29 日生效）	鼓励合成核酸提供者实施全面、可扩展且可验证的筛查机制，以防止这些核苷酸被滥用。基于 HHS 早期的指导，要求接受联邦研究和发现资金的机构仅从实施这些最优方法的提供者处采购合成核酸
	NIH 关于涉及重组或合成核酸分子研究的指南	为接受 NIH 资助，开展“涉及重组或合成核酸分子的基础和临床研究，包括创建和使用包含重组或合成核酸分子的生物体和病毒”的机构，制定安全实践和 g 隔离程序

政策重点与目标	具体法规、政策或指南	简要描述
公共获取与数据管理及共享： 力求在最大程度上实现研究成果的妥善管理与共享，同时确保研究人员考虑如何保护人类研究参与者的隐私、权利及数据保密性。通过数据共享提高数据的可得性，有助于更准确地开发和使用人工智能模型。这些政策有助于确保研究人员妥善管理人工智能模型使用或生成的数据。HH 各业务部门制定了一系列完善的政策，以规范数据共享，其中包括但不限于 NI 基因组数据共享政策、NIH 公共获取政策以及 NIH 数据管理与共享政策。	公共访问政策	2022 年 8 月，科学技术政策办公室发布了一份公共访问备忘录，指示进行研究和开发支出的联邦机构，在 2025 年 12 月 31 日前，使所有经过同行评审的学术出版物可公开获取，且无访问限制或费用。此外，这些出版物所依据的所有科学数据在默认情况下，应在出版时免费且公开可获取。作为回应，HHS 的运营部门已更新和 / 或制定了公共访问政策，以满足这一要求（见 NIH 公共访问政策）
	NIH 数据管理与共享政策	规定了提交数据管理与共享计划并遵守 NIH 批准计划的要求。此外，NIH 各研究所、中心和办公室可要求在计划中包含额外或特定信息，以支持项目优先事项或扩大研究产生的科学数据的效用
	NIH 基因组数据共享政策	促进并推动使用 NIH 资金生成的大规模基因组数据的合理共享
许可、知识产权与技术转让	美国专利商标局关于人工智能的信息	提供与人工智能相关的专利资源以及有关人工智能知识产权政策的重要信息
	NIH 研究工具政策	期望获得资助的机构适当地传播、推广并允许对使用 NIH 资助开发的研究工具进行开放访问