U.S. Department of Health and Human Services: Strategic Plan for the Use of Artificial Intelligence in Health, Human Services, and Public Health

Strategic Plan

January 2025

United States Department of Health and Human Services

Contents

Acknowledgements and Disclaimer

Acknowledgements

HHS would like to thank the HHS AI Task Force, Steering Committee, working group co-leads, and writers for their contributions to this document and many hours of work above and beyond expectations. We are grateful to the many colleagues across HHS who provided thoughtful comments and engaged in developing the Strategic Plan. The Department would also like to share its enormous gratitude to the HHS AI Task Force Project Management Office for its leadership, coordination, and direction. Finally, HHS would like to acknowledge the broad set of stakeholders from across the sector who volunteered their time and perspectives to inform this Strategic Plan. We sincerely appreciate their constructive and critical contributions.

Disclaimer

The U.S. Department of Health and Human Services AI Strategic Plan does not modify or interpret any requirements under the Federal Food, Drug, and Cosmetic Act (FD&C Act), the Public Health Service Act, Food and Drug Administration (FDA) regulations, or others. Nor does this document constitute a guidance document within the meaning of Section 701(h) of the FD&C Act (21 USC. 371(h)), 21 CFR 10.115, or others. Further, this document does not establish any rights or obligations with respect to any member of the public.

Letter from the Deputy Secretary

Artificial intelligence (AI) has had an undeniable influence on health, human services, and public health. At the U.S. Department of Health and Human Services (HHS), we have been steadfast in our efforts to responsibly leverage AI to advance our mission across critical areas within HHS and throughout the sector.

At HHS, we are optimistic about the transformational potential of AI. These technologies hold an unparalleled ability to drive innovation by accelerating scientific breakthroughs, improving medical product safety and effectiveness, improving health outcomes through care delivery, increasing access to human services, and optimizing public health. However, our optimism is tempered with a deep sense of responsibility. We need to ensure that Americans are safeguarded from risks. Deployment and adoption of AI should benefit the American people, and we must hold stakeholders across the ecosystem accountable to achieve this goal. AI creates vast opportunities to improve our country's health and human services and better serve the American people, and HHS is already taking active steps to motivate the ethical and responsible use of AI so that it might improve people's lives.

We are excited to introduce the HHS AI Strategic Plan, which presents our approach to catalyze innovation, promote trustworthy AI development, democratize technologies and resources, and cultivate AI-empowered workforces and organization cultures. This Plan represents a significant milestone in our ongoing commitment to harness the power of AI to strengthen our nation's health and well-being.

We will continue to do our part at HHS, as detailed in this Plan, using available resources and levers to successfully deploy AI in health, human services, and public health. But success requires a whole-of-nation approach in partnership with industry, academia, patients, and countless others. We look to the rest of the ecosystem to join us in this mission.

Deputy Secretary Andrea Palm

Introduction

Purpose of the Plan

HHS's vision is to be a global leader in innovating and adopting responsible AI to achieve unparalleled advances in the health and well-being of all Americans. This HHS AI Strategic Plan (hereafter referred to as "Strategic Plan" or "Plan") provides a framework and roadmap to ensure that HHS fulfills its obligation to the Nation and pioneers the responsible use of AI to improve people's lives.

Over the past 50 years, the U.S. has undergone a profound change in the way individuals interact with digital technologies. AI holds tremendous promise and potential risk for health and human services. While AI has existed in some form since the mid-20th century, AI has become ubiquitous in recent years. This is also true for healthcare and will increasingly be true in human services delivery. New and emerging technologies are making it even more possible to predict diseases before symptoms appear, identify new drug targets with the potential to transform the standard of care, and more effectively match human services to people who need them the most. Given the trajectory of this technology, the potential for AI to fundamentally change health and human services will become even greater.

This Strategic Plan defines AI as outlined in section 5002(3) of the National Artificial Intelligence Initiative Act (15 U.S.C. 9401(3)): a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments.[1] Within this definition, AI can take many forms. In the healthcare sector, basic algorithms for performing tasks or solving problems have been widely used for decades. Advances in AI and machine learning (ML) capabilities are strengthening algorithms to go beyond narrow rules and become more predictive and general by analyzing or "learning" from available data to tailor model output more precisely to the characteristics of a specific individual or subpopulation.[2] Generative AI (GenAI), another type of AI, refers to technologies that analyze and learn from data to create ("generate") something new, such as data, text, images, sounds, or other types of information.[3]

Private sector interest and investment have fueled the rapid growth of AI and health AI (technologies used in health and human services) capabilities. AI technology, and in particular GenAI, has been growing rapidly over the last several years, with industry and academic settings producing over 60 notable ML models in 2023 alone.[4] Venture capital and private AI investments have increased substantially, accounting for over $55B in U.S. venture capital funding across industries in Q2 2024.[5] Investment in GenAI is projected to grow by up to 42% year over year through 2032, leading to a potential $1.3T market across industries.[6] For healthcare, start-ups have raised approximately $30B for AI over the last three years.[7] There is an additional need for investment in human services delivery to meet population needs (e.g., the World Health Organization [WHO] estimates that 3.5B people will require assistive technology by 2050, some of which may be enabled by AI).[8] To ensure the responsible use of AI, entities in the U.S. have seen an increase in the number of regulations that mention AI (25 in 2023, an increase

---

1 While this definition will be used as the basis of this Strategic Plan, alternative definitions may at times be used by HHS operating and staff divisions.

2 https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

3 https://www.fda.gov/science-research/artificial-intelligence-and-medical-products/fda-digital-health-and-artificial-intelligence-glossary-educational-resource HHS recognizes that there exist additional terms to describe AI (e.g., Foundational Model, Constitutional AI); for simplicity, this Plan primarily addresses traditional and GenAI.

4 https://aiindex.stanford.edu/report/

5 https://www.reuters.com/business/finance/ai-deals-lift-us-venture-capital-funding-highest-level-two-years-data-shows-2024-07-03/

6 https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

7 https://www.aha.org/aha-center-health-innovation-market-scan/2024-09-17-top-4-health-care-ai-investment-trends-watch

8 https://www.who.int/news-room/fact-sheets/detail/assistive-technology

of 56% from 2022).[9] Global, multinational, and other governmental entities around the world, including the United Nations, Group of Seven (G7), Organisation for Economic Co-operation and Development (OECD), and WHO, are likewise making guidance and strategies for the use of AI a priority.

AI has paved the way for an increasing array of scientific breakthroughs and, in some cases, may surpass human performance in tasks like image classification and visual reasoning.[10] AI also has the potential to dramatically improve the ability to identify relevant factors or predict outcomes. Furthermore, advances in AI and ML fuel the increased use of predictive models in the "back office" of health and human services, such as appointment scheduling and evidence and literature reviews for research.

AI has or will directly or indirectly affect every American's experience in health and human services. Therefore, AI development should take a "human-centered design" approach that ensures it focuses on providing real benefits for people who use or receive services supported by AI.[11] Some of the benefits—to be discussed in greater detail below—include:

- Accelerating scientific breakthroughs that could increase the quality and length of life
- Being used as part of a medical product or to develop medical products to improve safety and effectiveness
- Improving clinical outcomes and enhancing safety through innovations in healthcare delivery
- Improving equity and empowering participants through enhanced health and human services benefits delivery
- Forecasting risks and rapidly mobilizing resources to predict and respond to public health threats

Such potential does not come without risks. While AI can significantly improve many aspects of health and human services, it also presents possible risks that could lead to adverse impacts and outcomes. Depending on the data and model quality, AI can produce outputs that are incorrect or incomplete. When important decisions are made in part or in whole based on AI that is not accurate, people can be harmed or denied access, and resources can be misused. Further, researchers have found that AI can introduce and propagate bias, which may misclassify people's needs, negatively impact physical or mental health outcomes, and increase costs.[12, 13, 14] Responsible AI use should also ensure equitable access and beneficence, safeguard protected information, and involve appropriate consent where applicable, while also considering potential unintended negative impacts on society or the environment. It is important to note that these risks and considerations may manifest differently depending on the complexity of AI used (e.g., simple rule-based algorithms will carry different considerations than large language models [LLMs]). Regardless, AI use in health and human services must ensure and be accountable to appropriate human oversight, and AI should be viewed as a tool to support and inform efforts rather than the sole answer to problems in the existing landscape.

The federal government is working to assess the potential of AI while ensuring it is safe and equitable for all Americans. Maximizing opportunities and mitigating risks is core to HHS's long-standing mission: Enhance the health and well-being of all Americans by supporting effective health and human services and fostering sound, sustained advances in the sciences underlying medicine, public health, and social services. This mission is supported by and connected to the missions of our community partners, state, tribal, local, and territorial

[9] https://aiindex.stanford.edu/report/

[10] https://aiindex.stanford.edu/report/

[11] https://digital.gov/topics/human-centered-design/ Human-centered design refers to the philosophy and method that places people's experiences at the heart of service design.

[12] https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:24027:ed-1:v1:en Bias is defined as "systematic difference in treatment of certain objects, people, or groups in comparison to others, where treatment is any kind of action, including perception, observation, representation, prediction, or decision."

[13] https://pubmed.ncbi.nlm.nih.gov/31649194/ Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453.

[14] https://www.nimhd.nih.gov/resources/understanding-health-disparities/diversity-and-inclusion-in-clinical-trials.html

governments (STLTs), academia, and private sector partners. It requires HHS to continue aligning efforts and priorities to ensure quality and safety and address the Nation's evolving health and human service needs while finding a balance that encourages innovation and deploys the necessary guardrails. Similarly, it requires empowering end users (people, including patients, healthcare providers, and others) to shape how new technologies are responsibly integrated into their care and services by fostering collaboration throughout the innovation pipeline.

Recent advances in the capabilities, breadth of applicability, ease of use, and speed of adoption of AI also suggest it may affect health and human services faster and with greater impact than anticipated. HHS and its operating and staff divisions ("divisions") recognize the value and importance of operating at an enterprise level rather than just through isolated uses within specific units for standalone purposes. It is critical for HHS to set a clear strategy to ensure health and human services organizations are well positioned to take advantage of AI according to consistent principles and objectives. A clear strategy is also necessary to manage the portfolio of AI investments and ensure HHS builds upon synergies between its divisions.

HHS plays a crucial role in the sector: an investor in research and discovery, a health industry regulator, a catalyst for innovation in delivering health and human services, a provider of healthcare and human services delivery, and a protector of patient safety, rights, and privacy. As AI adoption varies across industries within HHS's purview, a responsible approach for development and adoption is required. HHS will use the existing regulatory structure to clarify guidance, offer new guidance where needed, and update oversight mechanisms as necessary in response to technological innovation. HHS will also seek new regulatory authorities where appropriate. While the evolving nature of AI will likely challenge regulatory paradigms, HHS will continue to use all available levers, including policy, funding, education and outreach, and others to meet the new technological reality and support stakeholders in the health and human services ecosystem.

Organization and Use of the Plan

Organization of the Strategic Plan

The Strategic Plan is specifically focused on articulating HHS's vision and goals for AI in health, human services, and public health. As one of the largest federal entities in the U.S. government, HHS divisions and activities cover the entire continuum of health and human services, from bench-side research to bedside care delivery; from drug discovery to surveillance; and from childhood poverty prevention to benefits for seniors and people with disabilities. Given this expansive purview, the Strategic Plan presents a unifying framework composed of seven domains to promote alignment across HHS policies, programs, and activities involving AI.

- *Primary domains* represent specific parts of the HHS value chain, including:
    - **Medical Research and Discovery:** Fundamental and pre-clinical research on the basic mechanisms of disease and life processes, their translation to medical innovations and clinical applications,[15] and their context to use in healthcare delivery as a whole
    - **Medical Product Development, Safety, and Effectiveness:** Drug, biological product, and medical device development, clinical trials and regulatory approval, manufacturing, and ongoing safety and effectiveness monitoring
    - **Healthcare Delivery:** Provision of healthcare services to individuals and populations to diagnose, treat, manage, and prevent diseases and promote health and well-being, as well as financing to support this delivery
    - **Human Services Delivery:** Provision of social services and assistance to individuals and families to meet basic needs for health, welfare, self-sufficiency, safety, and well-being
    - **Public Health:** Protection and improvement of the well-being of populations through preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society, organizations, public and private communities, and individuals
- *Additional domains* are functional areas that span primary domains and are required to implement the Strategic Plan:
    - **Cybersecurity and Critical Infrastructure Protection:** Protection and advancement of systems' security critical to health and human service functions to support the use of AI
    - **Internal Operations:** Policies, programs, and infrastructure used by HHS divisions for internal operations and functions enabling HHS to implement the Strategic Plan and accommodate rapid technological advancements

Within each primary domain, chapters follow a consistent structure:

- Introduction and context to AI in the domain
- Stakeholders engaged in the domain's AI value chain
- Opportunities for the application of AI
- Trends in AI
- Potential use cases and risks
- Action plan

This full version of the Plan is deliberately expansive to provide context and tangible examples for readers seeking a more detailed orientation. It includes more comprehensive discussion of the opportunities, trends, use cases and risks, including full, granular action plans. For a high-level perspective, please see the Overview that was developed to increase accessibility and utility to a broad set of readers.

HHS Use of the Strategic Plan

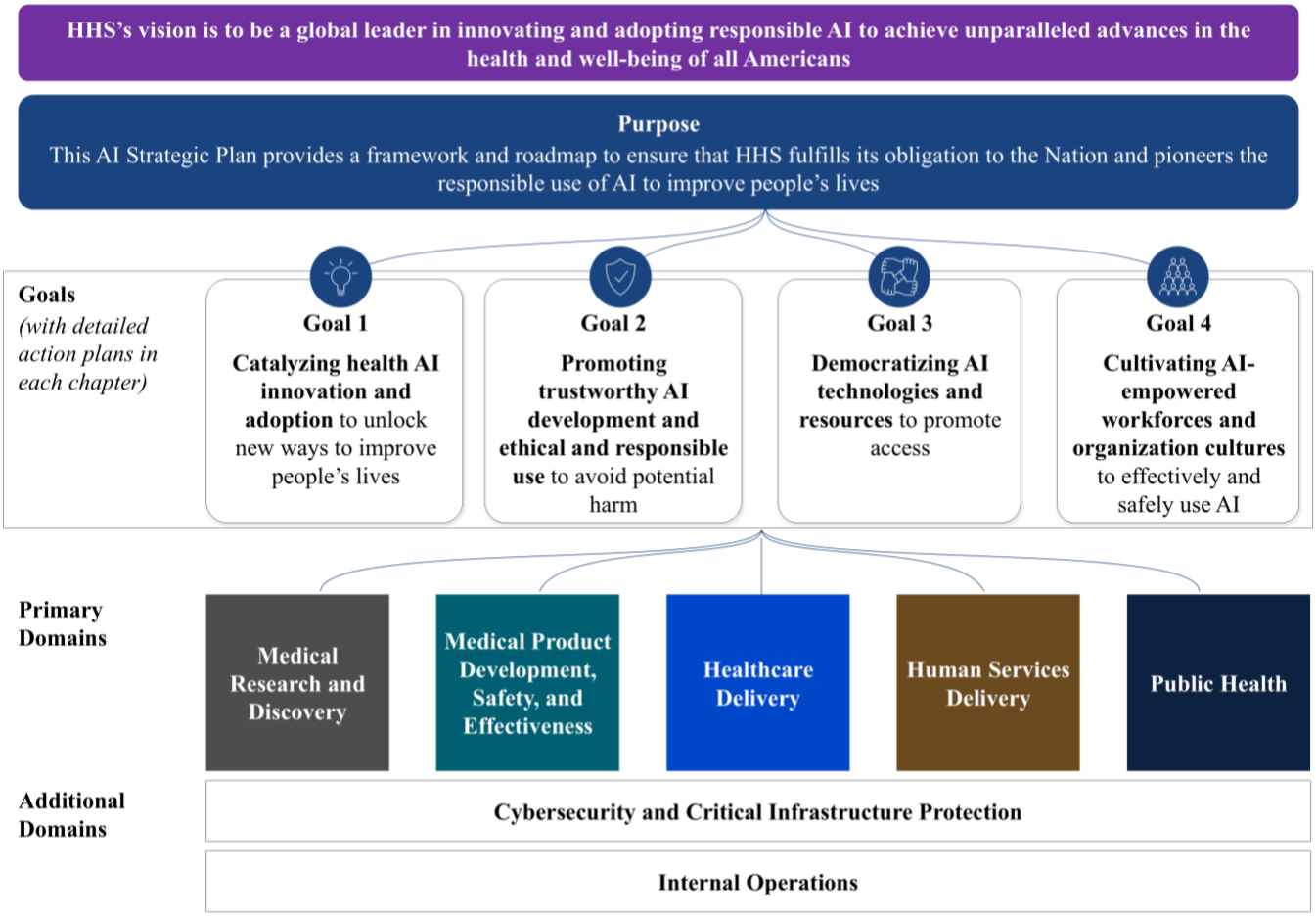
HHS's overarching objective is to set in motion a coordinated public-private approach to improving the quality, safety, efficiency, accessibility, equitability, and outcomes in health and human services through the innovative, safe, and responsible development and use of AI.

[15] HHS recognizes that the Medical Research and Discovery pipeline contains overlaps with Medical Product Development, Safety, and Effectiveness "development." However, for purposes of this Plan, AI use in pre-clinical research will be addressed in the Medical Research and Discovery chapter. Further steps will appear in the Medical Product Development, Safety, and Effectiveness chapter. Additionally, information on biosecurity will appear in the Medical Product Development, Safety, and Effectiveness chapter.

HHS will accomplish this objective by focusing on four key goals:

1. **Catalyzing health AI innovation and adoption** to unlock new ways to improve people's lives
2. **Promoting trustworthy AI development and ethical and responsible use** to avoid potential harm
3. **Democratizing AI technologies and resources** to promote access
4. **Cultivating AI-empowered workforces and organization cultures** to effectively and safely use AI

Exhibit 1: Goals and Structure of the Strategic Plan



As detailed in Exhibit 1, within each primary domain, chapters describe how HHS will focus on these four recurring goals through current and planned actions that will guide and support their execution. These actions will span a variety of levers available to HHS and its divisions, including regulations, policies and guidance, grants, funding programs, public education and outreach, and internal infrastructure, procurement, and operations. It is important to note that new policies are not the only way to support the responsible use of AI; existing approaches may be updated to address emerging concerns while ensuring that AI use remains compliant with current regulations (e.g., patient privacy). By orchestrating the use of these levers across its value chains, HHS aims to maximize coordination and strategically align its divisions and the rest of the health and human services ecosystem toward the achievement of HHS's strategic vision and the realization of the opportunities for AI to improve people's lives.

Opportunities for AI to Improve People's Lives

AI has the potential to improve people's lives and to support HHS's broader mission across areas. A few examples include:[16]

- **Accelerating scientific breakthroughs that could ultimately increase the quality and length of life:** Since 2000, the average timeline between Phase 1 clinical trials and regulatory approval has been approximately ten years, with even longer lead times for basic research and drug discovery.[17] Incorporating AI throughout the clinical discovery and development process offers tremendous hope in focusing on safe and effective targets, identifying populations and diseases for which products may be most effective, assessing the representativeness of the data and data models, and correcting for undersampling of populations, and more, ultimately shortening the development timeline and reducing overall costs.
- **Being used as part of a medical product or to develop medical products to improve safety and effectiveness:** AI can be used as part of a medical product or to develop safe and effective medical products. In particular, AI-enabled medical devices, such as over-the-counter hearing aids, have the potential to be used by patients, healthcare providers, and other end users to help augment care and improve outcomes.[18,19] Additionally, AI supports the ability to learn from data collected during clinical use which can help support improving medical product accuracy and performance over time,[20] potentially leading to improved accuracy and monitoring (e.g., lower misdiagnosis rates, higher ability to detect adverse effects early). Similarly, AI can be leveraged to help develop drugs and biological products (e.g., identifying targets, assessing biomarkers and endpoints).
- **Improving clinical outcomes and enhancing safety through innovations in healthcare delivery:** Medical errors, including incorrect and/or delayed diagnoses, may contribute to adverse outcomes.[21, 22] AI has the potential to accelerate diagnoses and head off safety events by rapidly processing expansive and disparate information, detecting patterns not always apparent to human observation, and directing clinicians to higher likelihood diagnoses and/or safety issues tailored to individual circumstances through clinical decision support and other tools. AI can also enhance care models and health services research to develop innovations that better enable clinicians, payers, and patients.
- **Improving equity and empowering patients and members of the public through improved health and human services benefits delivery:** Today, many individuals and communities face barriers to care given socioeconomic status, language, geographic location, and other factors.[23] AI has the potential to improve access to benefits and services for all individuals; for example, individuals for whom language is a barrier to receiving healthcare or human services may benefit from interpreter access through real-time, automated translation.[24] AI can also help individuals with disabilities perform simple or complex tasks, such as language technologies which can support individuals with speech impairments by optimizing speech patterns and turning them into fluent conversations.[25]
- **Forecasting risks and rapidly mobilizing resources to predict and respond to public health threats:** HHS has seen a significant uptick in the adoption of AI in response to public health crises such as the COVID-19 pandemic. At scale, AI has the potential to improve global infrastructure for predicting future

---

16 The chapters that follow detail the types of benefits specific to each domain.

17 https://www.mckinsey.com/industries/life-sciences/our-insights/generative-ai-in-the-pharmaceutical-industry-moving-from-hype-to-reality

18 https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

19 https://www.fda.gov/news-events/press-announcements/fda-authorizes-first-over-counter-hearing-aid-software

20 https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

21 https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/2813854

22 https://patientsafetyj.com/article/116529-patient-safety-trends-in-2023-an-analysis-of-287-997-serious-events-and-incidents-from-the-nation-s-largest-event-reporting-database

23 https://www.cdc.gov/health-equity/what-is/index.html

24 https://pubmed.ncbi.nlm.nih.gov/37904073/ Bakdash, L., Abid, A., Gourisankar, A., Henry, T. L. Chatting Beyond ChatGPT: Advancing Equity Through AI-Driven Language Interpretation. *J GEN INTERN MED* 39, 492–495 (2024)

25 https://www.forbes.com/councils/forbesbusinesscouncil/2023/06/16/empowering-individuals-with-disabilities-through-ai-technology/

disease outbreaks, enabling public health teams to develop effective countermeasures at scale prior to the first incidence of disease in new geographies. AI can be leveraged to improve public health through other means, such as identification of factors likely to impact health and human services (e.g., predicting natural disasters before they occur, which may reduce impact).

Promoting Ethical and Responsible Use of AI

The use of AI also carries several inherent challenges and risks. HHS is committed to developing, sharing, and promoting trustworthy AI that improves health and wellness outcomes. In support of this commitment, HHS is identifying existing practices to ensure trustworthy AI and addressing inconsistencies across domains. While it is not in the scope of this Plan to present a comprehensive approach to ethical and responsible use of health AI for every potential use case, HHS lays out overall considerations in this Plan that apply across the ecosystem. HHS expects all organizations to maximally promote ethical and responsible use of AI. Stakeholders should collectively work toward mitigating risks of inadvertent harms, such as falsely identifying patient conditions, breaching confidentiality of patient information (either directly or through reidentification of encrypted and/or deidentified patient data), misdirecting use of resources (particularly during public health emergencies), unintentionally developing potentially harmful medical products, or negatively contributing to social or environmental impacts. Stakeholders should also promote equity by reducing biases and increasing access for populations (e.g., geographic communities, persons with disabilities).

HHS will build on existing risk management and governance frameworks such as the National Institute of Standards and Technology (NIST) AI Risk Management Framework and Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology (hereafter "ASTP" or "ASTP/ONC") Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule (89 FR 1192). The NIST Framework asserts that holistic AI risk management requires risk mapping, measurement, and management to inform actions and governance. The HTI-1 Final Rule lays out a risk mapping approach for transparency of key information to assess benefits and risks of AI. Both NIST and certain policies finalized in the HTI-1 Final Rule are informed by the FAVES principles (fair, appropriate, valid, effective, and safe).

FAVES principles[26]
Fair: Model outcomes do **not exhibit prejudice or favoritism** toward an individual or group based on their inherent or acquired characteristics.
Appropriate: Model and process outputs are well matched to produce results **appropriate for specific contexts and populations** to which they are applied.
Valid: Model and process outputs have been **shown to estimate targeted values accurately** and as expected in both internal and external data.
Effective: Model outcomes have **demonstrated benefits in real-world** conditions.
Safe: Model outcomes are free from any known unacceptable risks, and the probable benefits outweigh any probable risks.

[26] https://www.healthit.gov/sites/default/files/2023-12/Health_Sector_AI_Commitments_FINAL_120923.pdf

FAVES is not an exhaustive list of all risk areas that can be considered, but its principles provide a foundation upon which AI development and use may be evaluated by describing the broad characteristics of high-quality AI within the context of health and human services.[27] Chapters of this Plan will discuss risks and mitigation strategies to ensure safe and trustworthy use. As AI advances rapidly, HHS will continue to revisit principles and engage stakeholders to respond to the challenges of AI. All individuals share responsibility to monitor for risks and support FAVES models and the use of AI.

Applicability to State, Tribal, Local, and Territorial Health and Human Services Organizations

In many cases, AI is deployed in individual STLTs as well as community-based organizations (CBOs). HHS recognizes that each organization has unique needs based on patient and population health factors and that, in some situations, organizations have differing responsibilities (e.g., some STLTs and CBOs provide direct services, whereas others do not). HHS will maintain a flexible approach that supports innovation while ensuring safe and responsible development and use. In this way, HHS and industry partners can learn from STLT and other entities as they increase their use of AI and identify new ways of improving health and human services. Relevant entities and potential actions are discussed in more detail in the domain-specific chapters.

In April 2024, HHS published a plan for promoting the responsible use of AI in automated and algorithmic systems by STLT governments in the administration of public benefits.[28] In this plan, HHS provides recommendations to STLTs on how they should choose, procure, design, govern, and manage AI in the administration of public benefits and services. The April 2024 plan also outlines HHS's plans to support STLTs in developing their own policies and practices for using AI in automated and algorithmic systems for public benefits programs and services. HHS maintains alignment with those recommendations in this strategy and describes additional priorities to support and enable STLT's safe and responsible development and use of AI.

HHS Roles and Responsibilities Relevant to AI

In alignment with the potential for AI to enhance the health and well-being of all Americans, HHS set up the Office of the Chief Artificial Intelligence Officer and established the role of the Chief AI Officer (CAIO) in March 2021. Located with ASTP, the primary functions of the CAIO are to drive implementation of the Strategic Plan, oversee the HHS AI governance structure, coordinate HHS's response to federal AI mandates, and foster AI-related collaboration. The CAIO has a vital role at HHS and within the federal government to maintain American leadership in AI. Fulfilling this commitment to AI within a department as vast and far-reaching as HHS requires coordination across divisions and department-wide alignment of responsible AI principles and resources. The CAIO will serve as this coordinating function, aligning the different divisions' diverse capabilities to advance the Strategic Plan. The CAIO will also monitor how cross-collaboration between divisions can create new opportunities for AI in health and human services, filling in gaps that a more diffuse strategy may miss. ASTP more broadly will also play a role in cross-HHS coordination of AI implementation and adoption.

[27] Risk of individual AI use cases or processes may need to be assessed along dimensions not included in the FAVES framework.

[28] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

HHS divisions below will play multiple roles in assessing opportunities for AI. Below is a brief description of each operating division and its key AI activities:

- **Administration for Children and Families (ACF):** Administers over 60 programs that provide benefits and services to support families and children, including promoting economic and social well-being. ACF's role in the HHS AI Strategic Plan will focus on ensuring effective and equitable delivery of human services to children and families.
- **Administration for Community Living (ACL):** Supports programs for populations with complex needs, particularly older adults and people with disabilities, and administers various programs, including nutrition services, elder support services, and elder rights programs. ACL's role in the HHS AI Strategic Plan will focus on ensuring effective and equitable delivery of human services to individuals with complex needs.
- **Agency for Healthcare Research and Quality (AHRQ):** Provides funding and programs to enhance quality, accessibility, equity, affordability, and safety in healthcare, including improvements in primary care and assistance in access to social welfare and public health services; management and oversight of the Patient Safety Organization program; award of investigator-initiated health services research funding inclusive of digital healthcare research, such as health AI and clinical decision support; and execution of national expenditure surveys capturing utilization, expenditures, and sources of payment and health insurance coverage. AHRQ's role in the HHS AI Strategic Plan will focus on promoting and conducting research on the adoption of safe AI and appropriate use in workflows to enable high-quality care.
- **Advanced Research Projects Agency for Health (ARPA-H):** Advances high-potential, high-impact biomedical and health research that cannot be readily accomplished through traditional research or commercial activities. ARPA-H's role in the HHS AI Strategic Plan will focus on issuing awards to catalyze cutting-edge research.
- **Administration for Strategic Preparedness and Response (ASPR):** Leads the nation's medical and public health preparedness for, response to, and recovery from disasters and other public health emergencies and collaborates with healthcare and public health stakeholders (e.g., STLTs and hospitals) and others to improve the country's readiness and response. ASPR's role in the HHS AI Strategic Plan will focus on coordinating the use of AI in public health emergencies (in collaboration with the Centers for Disease Control and Prevention [CDC] and other stakeholders).
- **Centers for Disease Control and Prevention (CDC):** Detects and responds to new and emerging health threats, conducts research, issues guidance, and designs programs that address the Nation's largest health problems, promote healthy and safe behaviors, communities, and environments, and train the public health workforce. CDC's role in the HHS AI Strategic Plan will focus on researching the efficacy of AI in disease prevention and implementing AI in public health efforts.
- **Centers for Medicare & Medicaid Services (CMS):** Administers the Medicare program, the federal portion of the Medicaid and CHIP programs, and the Health Insurance Marketplace®,[29] which together provide health coverage to approximately 50% of Americans. Additionally, CMS approves and oversees program waivers and demonstrations, develops and tests healthcare payment and service delivery models, develops health and safety standards for providers of healthcare services, implements quality initiatives, and promotes the adoption and use of health information technology, among other responsibilities. CMS's role in the HHS AI Strategic Plan will focus on determination of coverage for AI-enabled healthcare services as appropriate (using payment and regulatory policy to ensure trustworthy, responsible use of AI by payers and providers), oversight and certification of state information technology systems and data collection standards, and the provision of technical assistance to providers, states, and other stakeholders.
- **Food and Drug Administration (FDA):** Regulates medical products (including drugs, biological products, and medical devices) by evaluating their safety and effectiveness before and after marketing. FDA also advances public health by, among other things, fostering innovations that can help accelerate patient access to safe, effective, and innovative medical products. FDA also has the responsibility in maintaining the safety

[29] Health Insurance Marketplace® is a registered service mark of the U.S. Department of Health and Human Services.

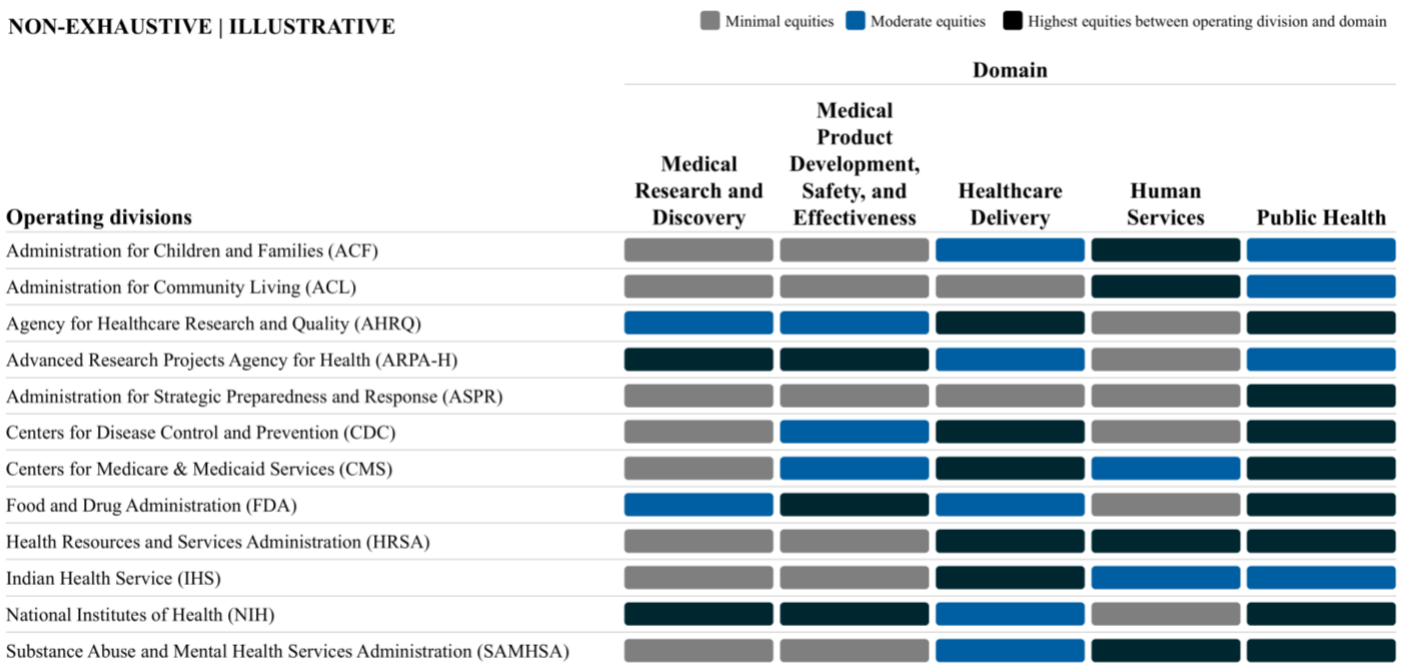
of our nation's food supply (human and animal), cosmetics, and products that emit radiation. In addition, FDA regulates the manufacturing, marketing, and distribution of tobacco products to protect public health. FDA's role in HHS's AI strategy will be focused on developing risk-based approaches to regulatory oversight of AI-enabled medical products and the AI used to develop medical products, issuing guidance for industry, and strengthening regulatory cooperation with international regulators.

- **Health Resources and Services Administration (HRSA):** Provides equitable healthcare to the nation's highest-need communities, including through programs that support people with low incomes, people with HIV, pregnant women, children, parents, rural communities, transplant patients, and the health workforce. This includes more than 31 million people cared for at HRSA-supported health centers, more than 58 million pregnant women, infants, and children, more than 560,000 people with HIV, more than 1,900 rural counties and municipalities across the country, and nearly 22,000 healthcare providers through loan repayment and scholarship programs. HRSA's role in the HHS AI Strategic Plan will focus on ensuring the equitable use of AI to benefit underserved communities and educating and training future generations of healthcare professionals.
- **Indian Health Service (IHS):** Provides primary and acute care for tribal nations and communities, representing approximately 2.8 million American Indians and Alaska Natives through a network of more than 600 hospitals, clinics, and health stations on or near Indian reservations. IHS's role in the HHS AI Strategic Plan will focus on implementing AI in healthcare delivery within these populations and ensuring the applicability of AI guidance to relevant STLTs.
- **National Institutes of Health (NIH):** Conducts and funds biomedical research and provides leadership and direction for programs designed to improve the Nation's health. NIH's role in the HHS AI Strategic Plan will focus on conducting and funding research to advance AI in biomedical, behavioral, and health research, developing and evaluating necessary standards, supporting the development of best practices for the training of AI models, developing and training AI workforce, and promoting the responsible use of AI.
- **Substance Abuse and Mental Health Services Administration (SAMHSA):** Leads efforts to reduce the impact of mental and substance use disorders on individuals, families, and communities. SAMHSA provides funding, guidance, and resources to support prevention, treatment, and recovery services, ensuring equitable access to care. SAMHSA's role in the HHS AI Strategic Plan will focus on providing grant funding and guidance to STLT communities and collecting, analyzing, and distributing behavioral health data to evaluate programs, improve policies, and raise awareness of resources on prevention, harm reduction, treatment, and recovery. SAMHSA will additionally support the adoption of AI by behavioral health clinicians and health systems.

HHS divisions have many areas of complementary and interdependent responsibilities. While operating divisions may span multiple areas, the following schematic depicts a general overview of division equities in each domain:

Exhibit 2: Overview of Equities of HHS Operating Divisions

Note: This schematic directionally indicates which divisions engage in which domains, necessitating coordination and collaboration. It is not meant to be an exhaustive indication of each division's equities, and divisions may play roles across domains in varied ways.



In addition to the operating divisions listed above, HHS staff divisions will play a large role in ensuring the success of the Strategic Plan. For example, ASTP oversees the adoption of data and technology standards for the access, exchange, and use of clinical information in healthcare, public health, and human services. It also guides the regulation of health information technology (e.g., electronic health records) in various federal programs and supports interoperability for government and industry constituents. ASTP's role will focus on cross-HHS policy and coordination of AI implementation and adoption. The Office for Civil Rights (OCR) enforces federal civil rights laws (e.g., Section 1557 Final Rule), conscience and religious freedom laws, the Health Insurance Portability and Accountability Act (HIPAA) Privacy, Security, and Breach Notification Rules, and the Patient Safety Act and Rule, which together protect fundamental rights of nondiscrimination, conscience, religious freedom, and health information privacy. OCR's role will be to provide education on protecting individuals' rights throughout AI development and use. The Office of the Assistant Secretary for Planning and Evaluation, the Office of Global Affairs, and the Office of the Assistant Secretary for Health have additional equities. The Office of the Chief Information Officer will also have a notable role in supporting internal uses of AI at HHS. This is not an exhaustive list of all HHS staff divisions or the entirety of work each will perform, but a way to highlight the extensive workstreams and responsibilities across the Department and articulate the importance of coordination. Individual as well as collaborative efforts across all HHS divisions will be critical in supporting this Strategic Plan.

Action Plan Summary

The following chapters will articulate existing and planned activities that support these goals. These actions are organized into themes that detail HHS's aspirations for the future of AI as articulated in the table below.[30]

Key goals that actions support	Themes of actions across chapters *(non-exhaustive, detailed Action Plans appear in each chapter)*
1. Catalyzing health AI innovation and adoption to unlock new ways to improve people's lives	• Expanding breadth of AI use across the value chains in each domain • Modernizing infrastructure to implement AI and support adoption • Enhancing collaboration and public-private partnerships to promote AI adoption • Clarifying regulatory oversight and coverage/payment determinator processes for AI • Supporting gathering evidence on outcomes (e.g., efficacy, safety) of AI interventions and best practices
2. Promoting trustworthy AI development and ethical and responsible use to avoid potential harm	• Building and disseminating evidence that supports mitigating risks to equity, biosecurity, data security, and privacy • Setting clear standards that guide the use of federal resources in the context of trustworthy AI use • Supporting organizational governance for risk management of AI • Refining regulatory frameworks to address adaptive AI technologies • Promoting external evaluation, monitoring, and transparency reporting and fostering other mechanisms for quality assurance of health AI
3. Democratizing AI technologies and resources to promote access	• Increasing access to responsibly curated data and infrastructure, including providing support for organizations where appropriate • Supporting information-sharing mechanisms to disseminate standards, best practices, and foster collaboration to improve access • Developing user-friendly, customizable, and open-source AI tools • Enhancing capabilities of STLTs and other community organizations, including providing resources or other mechanisms where appropriate
4. Cultivating AI-empowered workforces and organization cultures to effectively and safely use AI	• Improving training in governance and management of AI • Developing and retaining a robust AI talent pipeline • Equipping professionals with access to resources and research to support their respective health and human services organizations • Using AI to mitigate labor workforce shortages and address burnout and attrition

HHS's vision is to be a global leader in the innovative and responsible development and adoption of AI to achieve unparalleled advances in the health and well-being of all Americans. The following chapters of this Strategic Plan detail specific actions to achieve that vision.

[30] Some themes and actions may be repeated across chapters when they apply across domains

1 Medical Research and Discovery

1.1 Introduction and Context

Medical research and discovery are fundamental to advancing health by driving the development of innovative drugs,[31] biological products,[32] medical devices,[33] including some software-based behavioral interventions,[34] and other tools that improve individuals' and communities' health outcomes and access to quality care.[35]

This chapter of the Plan will focus on the research and discovery of medical products[36] and the research and discovery of AI technologies that can be leveraged in biomedicine. The next stages of the medical product life cycle, including clinical trials, as well as research in other fields, such as health systems, human services delivery, and public health, will be discussed in other chapters and are not in the scope of this chapter.[37]

In recent years, medical technology and pharmaceutical companies, academic and research institutions, and other organizations have increasingly leveraged AI to bolster their medical research and discovery activities and create AI-driven tools, but the full opportunity of existing AI technology is not captured today. While further advancements could unlock additional benefits, action is required to catalyze safe and responsible uptake of AI that more fully realizes the potential of AI in medical research and discovery settings. Accordingly, this chapter of the Plan explains the industry trends, AI use cases and risks, and actions that HHS could pursue to help safely activate AI adoption in medical research and discovery. HHS provides high-level context on medical research and discovery and an overview of AI in the space, including the stakeholders involved and key opportunities for AI uptake.

Medical research and discovery provide the data and the confidence to evaluate diagnostics, therapeutics, treatments, vaccines, technologies, and other tools in humans for the diagnosis, prevention, mitigation, and treatment of disease. At a high level, they can be described in a value chain that includes three phases: basic research, discovery (which can vary between different types of medical products), and pre-clinical studies. See Section 1.5 "Potential Use Cases and Risks for AI in Medical Research and Discovery" below for a detailed discussion of this value chain and its constituent phases.

Across all aspects of medical research and discovery, HHS plays an active role in spurring activity and promoting safety and quality. Nearly 83% of NIH's funding is awarded for extramural research and research support;[38] furthermore, NIH follows the HHS Common Rule[39] and has its own policies to ensure the safety of human research subjects, maintain data security and quality, and provide additional protections for vulnerable

---

31 See Appendix A: "Glossary of terms" for the definition of "drug" used in this Plan.

32 See Appendix A: "Glossary of terms" for the definition of "biological product" used in this Plan.

33 See Appendix A: "Glossary of terms" for the definition of "medical device" used in this Plan.

34 Note that some software-based behavioral interventions are medical devices under FDA's statute, whereas others, such as those software functions that are "intended for maintaining or encouraging a healthy lifestyle" and are "unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition," are not. See sections 201(h) and 520(o)(1)(B) of the FD&C Act.

35 https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-research-and-development

36 Drugs, biological products, and medical devices in this Plan are referred to as "medical products" when discussed collectively. See Appendix A: "Glossary of terms" for the definition of "medical products" used in this Plan for additional details.

37 Note that research pertaining to health systems, care delivery, and non-device behavioral interventions will be discussed in the "Healthcare Delivery" chapter; research pertaining to human services delivery will be discussed in the "Human Services Delivery" chapter; and research pertaining to public health will be discussed in the "Public Health" chapter. Furthermore, where relevant, clinical trials will be discussed in the "Medical Product Development, Safety, and Effectiveness" chapter and are not in the scope of this chapter.

38 https://www.nih.gov/about-nih/what-we-do/budget

39 https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html

communities participating in research.[40] Additional divisions also play transformative roles: in FY 2023, ARPA-H and AHRQ had budgets of $1.5B and $374M, respectively, to advance groundbreaking innovation in biomedicine and health.[41, 42, 43] In addition, FDA regulates scientific studies that are designed to develop evidence to support the safety and effectiveness of investigational drugs (human and animal), biological products, and medical devices.[44, 45] Though this summarizes a few of HHS divisions' roles in medical research and discovery, many more engage in the space in other ways. As AI becomes increasingly used in medical research and discovery, HHS and its core engaged divisions will facilitate the safe and impactful uptake of equitable AI technologies across the ecosystem.

1.1.1 Action Plan Summary

Later in this chapter, HHS articulates proposed actions to advance its four goals for the responsible use of AI in the sector. Below is a summary of the themes of actions within each goal. For full details of proposed actions please see section 1.6 Action Plan.

Key goals that actions support	**Themes of proposed actions *(not exhaustive, see 1.6 Action Plan for more details)***
1. Catalyzing health AI innovation and adoption	• Expanding the breadth of medical research and discovery AI use across disease areas and steps of the value chain • Enhancing coordination across geographies to harness AI to improve medical research and discovery • Fostering AI-ready data standards and datasets to bolster their usability for AI-empowered medical research and discovery
2. Promoting trustworthy AI development and ethical and responsible use	• Building and disseminating evidence to mitigate biosecurity, data security, privacy, and data collection risks • Setting clear guidelines for safe and trustworthy AI use in medical research and discovery and the distribution and use of federal resources • Enabling safe and responsible organizational governance of AI risk management and transparency
3. Democratizing AI technologies and resources	• Fostering intentional public engagement and public-private action to enhance sharing of best practices among all stakeholders • Increasing accessibility to responsibly curated AI-ready data, models and algorithms, and tooling and infrastructure for all
4. Cultivating AI-empowered workforces and organization cultures	• Improving training in governance and management of AI in medical research and discovery • Developing and retaining a robust AI talent pipeline in medical research and discovery

[40] https://grants.nih.gov/policy-and-compliance/policy-topics/human-subjects/policies-and-regulations

[41]. https://arpa-h.gov/sites/default/files/2023-10/FY_2023_NIH_ARPA-H_Operating_Plan.pdf

[42] https://www.ahrq.gov/news/blog/ahrqviews/ahrq-2024-proposed-budget.html

[43] https://arpa-h.gov/about/faqs

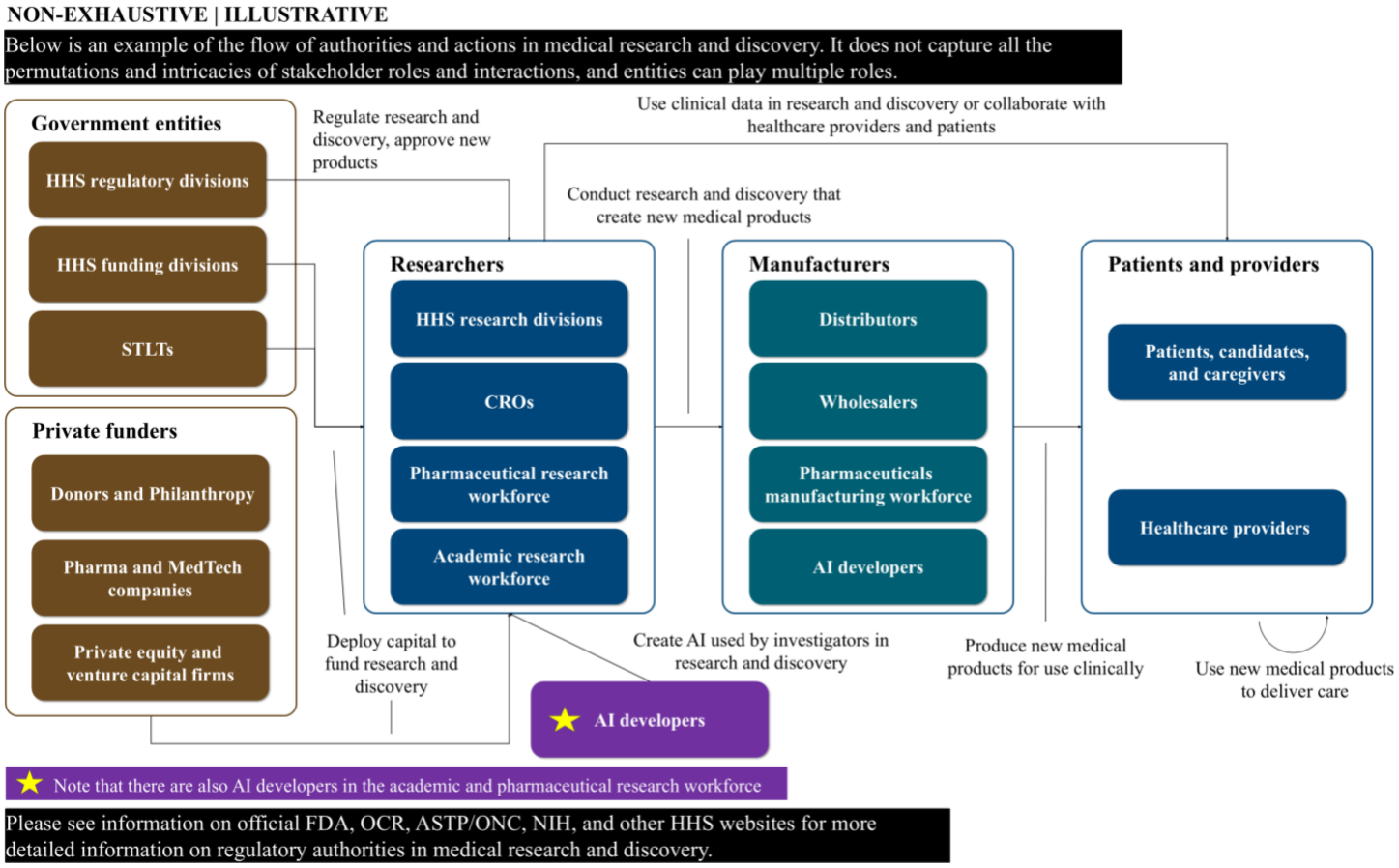
[44] https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process

[45] Note that FDA also oversees clinical research to ensure trials are designed, conducted, analyzed, and reported according to federal law and FDA's good clinical practice (GCP) regulations,[45] and after research, discovery, and any clinical trials are completed, the FDA reviews the data and information provided for marketing authorization and monitors authorized products postmarket to help ensure they remain safe and effective (see "Medical Product Development, Safety, and Effectiveness" for additional details).

1.2 Stakeholders Engaged in the Medical Research and Discovery AI Value Chain

Medical research and discovery must ultimately meet the needs of current and future patients and their caregivers; therefore, corresponding AI use should advance research and eventual technologies that meet these needs. In addition to patients and medical providers, several key stakeholders engage with AI in medical research and discovery, ranging from developers of medical products to distributors, providers, payers, researchers, and many others. The Action Plan section at the end of this chapter includes approaches to engage these stakeholders to advance innovation while mitigating risks. Below is an illustrative diagram of example flows between stakeholders and a bulleted list with additional details on medical research and discovery stakeholders. Please note that neither the diagram nor the list captures all possible stakeholder roles and interactions. Please refer to other HHS documents for additional regulatory guidance and authority details.

Exhibit 3: Stakeholders Engaged in Medical Research and Discovery



Stakeholders (including partners) include:

- **HHS operating divisions (non-exhaustive):**[46] Divisions involved in AI for medical research and discovery include:
    - **NIH:** Supports biomedical and behavioral research within the U.S. and abroad, conducts research in its own laboratories and clinics, trains promising young researchers, and promotes collecting and sharing biomedical knowledge. In recent years, these activities increasingly included AI related to medical research and discovery (e.g., making data available, catalyzing data science and AI

[46] https://www.hhs.gov/about/agencies/hhs-agencies-and-offices/index.html

opportunities in biomedical research and discovery, increasing diversity in AI model development, and developing and implementing AI across biomedical research domains).[47]
    - **ARPA-H:** Accelerates better health outcomes for everyone by supporting the development of high-impact solutions to society's most challenging health problems, including those leveraging AI (e.g., using AI to speed up the discovery and development of antibiotics).[48]
    - **FDA:** Helps ensure that human and animal drugs, biological products, and medical devices are safe and effective for their intended uses and that electronic products that emit radiation are safe. As AI becomes a more prominent aspect of medical research and discovery, the FDA will continue to play a role in regulating products and supporting stakeholders.
    - **AHRQ:** Focuses on improving the quality, safety, efficiency, and effectiveness of healthcare for all Americans through research, technology assessments, and work on dissemination and implementation. AHRQ will focus on promoting and conducting research on the safe adoption of AI that enables high-quality care, disseminating actionable, evidence-based AI knowledge, and provisioning evidence required for coverage decisions.
- **Other federal agencies:** HHS also works closely with many other federal departments, such as the National Science Foundation (NSF) and the Department of Energy (DOE).
- **Patients, research participants, caregivers, and related advocacy groups (including residents and communities):** Historically, considered the recipients or administrators of diagnostics, therapeutics, treatments, vaccines, technologies, and other tools designed by and/or embedded with various types of AI. Though patient centricity is not novel, empowered patients may now also utilize AI to understand their personal health status better and advocate for their own care; they can be included in the research and discovery process (e.g., as collaborators in the early planning phases of a study).[49]
- **Academic, non-profit, and other research workforce:** Investigators developing evidence to drive forward the leading edge of biomedical knowledge, engineers designing and generating medical devices for application in the clinic, and subject matter experts that develop AI, apply AI in research workflows, and/or integrate AI into the product development life cycle. They are among the primary users of AI in medical research and discovery.
- **Pharmaceutical, biotechnology, and medical device industry research workforce:** Responsible for the design, development, and production of diagnostics, therapeutics, treatments, vaccines, technologies, and other tools for commercial use in healthcare delivery, including researchers and subject matter experts integrating AI into research workflows and product design. They are among the primary users of AI in medical research and discovery.
- **Healthcare providers:**[50] Hospitals, clinics, and healthcare professionals who utilize medical products are often looped into medical research and discovery to provide clinical perspectives. Additionally, providers can serve as "humans in the loop" for medical research and discovery value chains.
- **State, tribal, local, and territorial governments (STLTs):** Regulatory agencies outside the federal government. While medical products are under the regulatory control of the FDA, the practice of medicine generally is under the jurisdiction of STLTs. Additionally, STLTs can fund medical research and discovery activities.[51]
- **Distributors and wholesalers:** Facilitate the distribution of medical products—which may have been researched and discovered by leveraging AI—to healthcare providers.
- **Contract research organizations (CROs):** Provide outsourced research services, potentially more concentrated in clinical development, which is elaborated on in the Medical Product Development, Safety,

[47] https://datascience.nih.gov/artificial-intelligence

[48] https://arpa-h.gov/news-and-events/arpa-h-project-accelerate-discovery-and-development-new-antibiotics-using

[49] https://heal.nih.gov/resources/engagement/understanding-pce

[50] Note that healthcare providers do not just adopt medical products but also implement evidence generated from research into care delivery, as well as healthcare delivery models and practices. They are also often research sites or research participants. See the "Healthcare Delivery" chapter for additional information.

[51] https://ncses.nsf.gov/surveys/state-government-research-development/2023

and Effectiveness chapter, and may develop or integrate AI into their medical research and discovery value chains or workflows.

- **Donors and private funders:** Non-profit donors, such as foundations and for-profit funders, such as private equity, venture capital, and other funding organizations, play a role in medical research and discovery and ongoing development by supporting funding for upstream research. These organizations may also support direct investment in aggregating datasets, developing platforms or AI tools, and using AI in the process.
- **AI-first technology developers:** Engineers and organizations who build the AI tools (e.g., protein-folding software), models, data infrastructure, and platforms (e.g., electronic health records) that can be used throughout the medical research and discovery value chain. Developers include AI-first biotechs, big tech, and domain-specific players.

HHS will engage stakeholders in the development or refinement of any funding mechanisms, policy guidelines, educational materials, or internal infrastructure relevant to AI in research and discovery to ensure HHS promotes equity in the access, understanding, and impact potential of these technologies. Furthermore, working closely with STLTs, particularly their regulatory bodies for health and human services, will allow this Plan to be aligned across levels of government and throughout geographies. Engaging stakeholders throughout the ecosystem will be critical to executing this work.

1.3 Opportunities for the Application of AI in Medical Research and Discovery

Responsible adoption and scaling of AI across the medical research and discovery value chain has the potential to improve health outcomes and access for Americans by:

1. **Bolstering the potential for basic research to derive novel biological insights that improve human health:** Not all medical research and discovery is directly "translational" (i.e., aiming to produce results immediately actionable in medical care). In fact, "basic" research (i.e., aiming to understand a phenomenon or mechanism more deeply) has historically led to some of the most impactful downstream impacts on human health (e.g., CRISPR).[52, 53] By leveraging AI to examine links between diseases and core pathological processes with data from clinical use (e.g., in longevity research), explore more hypotheses based on rapid analysis of very large volumes of data, screen images to augment human investigation, and generate insights at high speed, new basic research discoveries could not only proliferate but also be of higher quality than those arrived at without the support of AI.[54] Most importantly, this transformation could lead to better human health.
2. **Increasing accessibility to drive innovation and potentially reducing costs:** Emerging evidence suggests that leveraging AI across the medical research and discovery value chain presents a financial opportunity, up to $26B annually just for drugs[55] with potential additional value for devices. If realized, such efficiencies could lower barriers to conducting medical research and discovery and/or free up capital for reinvestment into further medical research and discovery activities. For example, medical research and discovery costs can be driven substantially by "wet lab" real estate, a space where physical biological and chemical samples can be tested, which may cost nearly double the asking rent of traditional office space per square foot.[56, 57] By leveraging AI to conduct some steps of medical research and discovery (e.g., protein folding modeling, simulations of biological interactions) *in silico*, the need for wet lab space could be reduced, which may

[52] https://www.nih.gov/news-events/gene-editing-digital-press-kit

[53] https://www.niaid.nih.gov/grants-contracts/basic-research-definition

[54] https://pmc.ncbi.nlm.nih.gov/articles/PMC10018490/

[55] https://itif.org/publications/2020/12/07/fact-week-artificial-intelligence-can-save-pharmaceutical-companies-almost/

[56] https://www.cbre.com/press-releases/net-absorption-of-lab-space-grew-nationally-in-the-second-quarter

[57] https://mktgdocs.cbre.com/2299/ebd1da98-2b86-4a75-b3ed-b050fb52d383-283656098/Q3_2024_U.S._Office_Figures_D3.pdf

lower costs required to engage in innovation. AI can allow institutions with lower access to capital (e.g., start-ups, non-profits, academic research organizations) to participate in innovation, increasing diversity in medical research and discovery that can lead to more breakthroughs. Furthermore, these potential reductions in cost could spur opportunities if reinvested. While costs and timelines vary from product to product, total development costs of some drugs, for example, can range from $300M to $4.5B each.[58] If the potential $26B annual financial opportunity is realized and reinvested, this could materially accelerate the availability of new innovations.

3. **Expanding the reach of medical research and discovery to meet unmet patient needs and support breakthrough innovations:** AI may foster breakthrough innovations and the development of novel medical products that address the health needs of patients who have been historically underserved. Research and discovery activity today may focus on potentially more profitable therapeutic areas (TA) rather than TAs with the most health need[59], given the significant cost and time associated with the research and discovery of a single medical product (see trend 2 in Section 1.4 below for more details). Leveraging AI to expand research and discovery beyond such "safe bet" targets or diseases and to increase pipeline activity on potentially under-researched TAs while pursuing breakthrough innovations across other TAs could transform outcomes and access for patients with underserved health needs. By leveling the field of targets or TAs "worth exploring," AI could also reduce bias in basic medical research and discovery.
4. **Accelerating the timeline to develop new products and potentially access care:** Currently, pre-clinical development for drugs, in particular, is estimated to take between six and ten years.[60] In recent years, however, leveraging AI in medical research and discovery has shown promise in corresponding use cases (e.g., from years for humans to determine protein structures to mere seconds).[61] If AI is successfully and responsibly adopted and scaled across the medical research and discovery value chain, these efficiencies could significantly reduce the time required to get medical products to patients, saving American lives, improving health outcomes, and more rapidly reaching underserved patients.[62] As the world leader in medical research and discovery, the U.S. could accelerate access globally as well.[63]

1.4 Trends of AI in Medical Research and Discovery

Adoption of AI in medical research and discovery is growing, following a few key trends:

1. **AI adoption is increasing yet inconsistent across the medical research and discovery value chain:** To date, uptake has focused more on deterministic activities in discovery, particularly in target identification and lead generation (e.g., predicting protein folding, molecular interactions, and cellular disease processes). Specifically, *in silico* design, manipulation, and exploration of biomolecules and designs of devices may have achieved more adoption of AI than use cases in basic research or pre-clinical studies (see the Potential Use Cases and Risks for AI in Medical Research and Discovery section below for examples of use case adoption across the value chain).[64]
2. **AI uptake is potentially concentrated on TAs with stronger market incentives:** Researchers face strong incentives, such as lucrative IP ownership, to focus medical research and discovery activities on profitable TAs that AI adoption does not necessarily address and may even exacerbate (e.g., more data leading to better models that are leveraged for further research and discovery on lucrative TAs).[65] AI investments may face similar incentives to focus on use cases related to exploring "high-confidence targets," which could

[58] https:/pmc.ncbi.nlm.nih.gov/articles/PMC11214120
[59] https://pmc.ncbi.nlm.nih.gov/articles/PMC3796018/
[60] https://pmc.ncbi.nlm.nih.gov/articles/PMC5725284/
[61] https://pmc.ncbi.nlm.nih.gov/articles/PMC11292590/
[62] https://allofus.nih.gov/news-events/research-highlights/all-of-us-artificial-intelligence-help-speed-up-search-for-promising-medicines
[63] https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-research-and-development
[64] https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2819343
[65] https://pmc.ncbi.nlm.nih.gov/articles/PMC3796018/

include a concentration on known, rather than novel, targets.[66] With the right interventions to overcome these structural incentives, however, AI could be leveraged toward less researched targets and TAs and achieve breakthrough innovations that meet unmet patient needs, which is a large opportunity as highlighted above in Section 1.3, opportunity 3 "expanding the reach of medical research and discovery to meet unmet patient needs and support breakthrough innovations."

3. **Medical research and discovery are extending beyond traditional laboratories:** While investigators use AI to expedite medical research and discovery, other players—such as technology companies—are also entering the research and discovery ecosystem with novel AI innovations. For example, defining the dynamic structure of proteins used to require crystallography, an arduous process through which proteins are crystalized with X-ray diffraction elucidating the position of their atoms, which required access to wet lab space. Recent investments led to the development of an open-source algorithm that can predict the structure of many proteins and how they fold and interact with other proteins and molecules in the body.[67] Some experiments can now be done significantly faster *in silico*. However, these first-pass results should still be validated through biological methods and/or have humans in the loop to ensure accuracy. While technology companies pursue solutions like these, pharmaceutical and medical technology companies are also building AI applications,[68] which can be leveraged to transform the quality of tasks across the medical research and discovery value chain and accelerate the time it takes to accomplish them.
4. **Data are fragmented, and infrastructure costs are rising:** Successful adoption of AI in medical research and discovery requires access to large amounts of high-quality training data, which are critical to the foundation of ML and other models.[69] Today, approximately 75% of scholarly documents, which contain data that could be leveraged in medical research and discovery AI models, is behind paywalls[70] (which may change as public access policies[71] are implemented). The potentially large quantities of data that could be very useful for medical research and discovery that do not exist in the "scholarly record" are fragmented and difficult to aggregate and curate (e.g., real-world data).[72] Furthermore, the specialized hardware and computing required to utilize AI can be expensive[73] and require high energy consumption. Entities with fewer resources to acquire this technology may be priced out, hindering equitable adoption and limiting innovation. These limitations will be compounded without equitable and safe access to data for AI in medical research and discovery.
5. **Agentic AI and other autonomous systems are potentially growing:** HHS is committed to using AI ethically and safely, including any potential adoption of agentic AI.[74] While currently nascent, agentic AI—systems with autonomous problem-solving and collaborative capabilities—is poised to become part of the lab to help augment researchers' activities across the medical research and discovery value chain. Unlike traditional AI, which follows programmed rules, agentic AI can independently or collaboratively analyze, decide, and act. Agentic AI could make medical research and discovery faster and, in turn, make breakthrough innovations available to patients sooner. HHS is already taking action to get ahead of this trend; for example, ARPA-H has released a request for information to understand agentic AI and set its corresponding strategic direction for medical research and discovery.[75]

---

66 https://pubmed.ncbi.nlm.nih.gov/37479540/

67 https://pmc.ncbi.nlm.nih.gov/articles/PMC11292590/

68 https://www.nature.com/articles/d41586-024-02842-3

69 https://aspe.hhs.gov/training-data-machine-learning-enhance-patient-centered-outcomes-research-pcor-data-infrastructure

70 https://pmc.ncbi.nlm.nih.gov/articles/PMC6825414/

71 https://sharing.nih.gov/public-access-policy

72 https://pmc.ncbi.nlm.nih.gov/articles/PMC6587701/

73 https://cloud.nih.gov/resources/guides/cloud-introduction/why-the-cloud/

74 https://arpa-h.gov/news-and-events/rfi-agentic-artificial-intelligence-systems

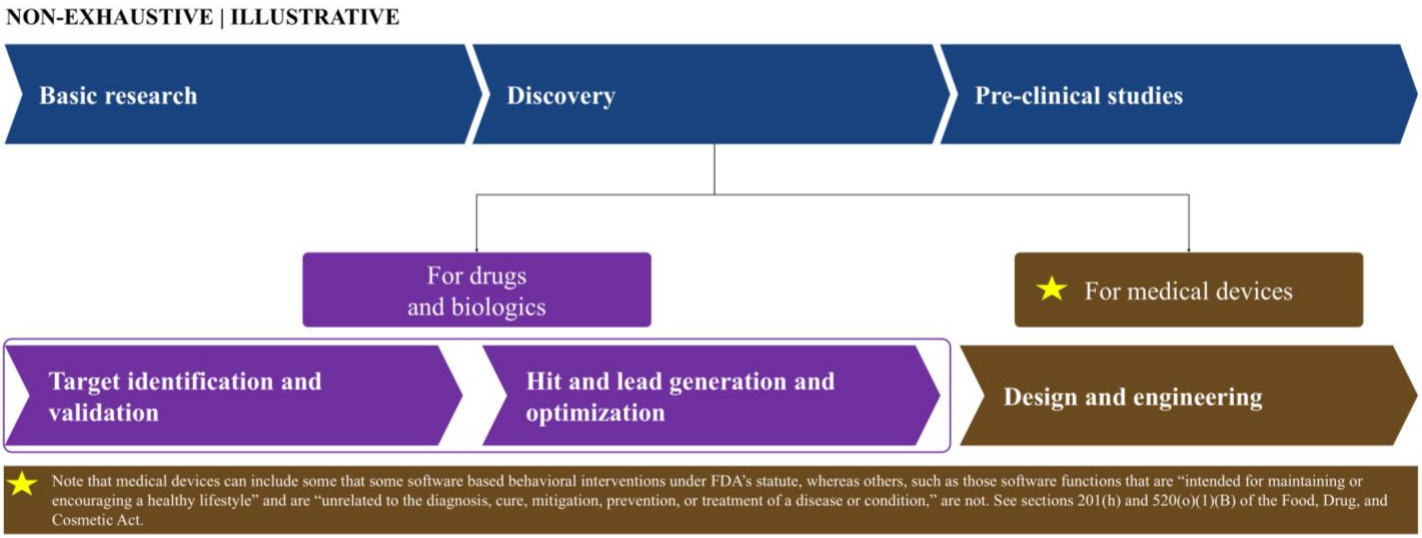
75 https://arpa-h.gov/news-and-events/rfi-agentic-artificial-intelligence-systems

1.5 Potential Use Cases and Risks for AI in Medical Research and Discovery

The Medical Research and Discovery Value Chain

In the U.S., medical research and discovery is a rigorous, multistep process aimed at bolstering knowledge of biology and ensuring the safety and efficacy of drugs, biological products, and medical devices before they reach the market. While there can be variation, in general, it forms a three-step value chain: (1) basic research, (2) discovery, which has different steps for different types of products, and (3) pre-clinical testing.[76] Clinical trials, where relevant, will be discussed in the Medical Product Development, Safety, and Effectiveness chapter and are not in the scope of this chapter. Similarly, research on health systems, care models, and behavioral interventions that are not medical devices is not in the scope of this chapter and is included in Healthcare Delivery. Also, this value chain of medical research and discovery activities can inform additional areas, such as public health, healthcare delivery, and human services delivery, in an iterative feedback loop.

Exhibit 4: Medical Research and Discovery Value Chain



1. **Basic research** involves scientific exploration that can reveal fundamental mechanisms of biology, disease, or behavior[77] to advance general knowledge or understanding of biological phenomena and observable facts, which are fundamental to advances in human health and one reason NIH funds basic research.[78] The small steps forward at the leading edge of a field can lead to new biomarkers or mechanisms of action for developers to target and give investigators and the public confidence in eventually testing new drugs, biological products, medical devices, technologies, and other tools with human research participants outside this step of the value chain.
2. **Discovery** is the scientific exploration to diagnose, treat, or cure disease, which can vary by type of medical product as described below:[79]
    a. **For drugs**[80] **and biological products**[81] (e.g., therapeutics, vaccines):
        i. **Target identification and validation** are important to the early stages of drug development, which generally relies on the initial identification of a suitable biological target for drug

[76] Note that the value chain for drugs and biological products versus medical products differs in the Discovery step, detailed below.
[77] https://ncats.nih.gov/about/about-translational-science/spectrum#basic-research
[78] https://grants.nih.gov/policy-and-compliance/policy-topics/clinical-trials/besh
[79] https://toolkit.ncats.nih.gov/module/discovery/
[80] See Appendix A: "Glossary of terms" for the definition of "drug" used in this Plan.
[81] See Appendix A: "Glossary of terms" for the definition of "biological product" used in this Plan.

candidates.[82] This includes finding the biological systems (e.g., neural circuits, endocrine, or immune pathways) that a therapeutic can regulate and ensuring that engagement of that target has a "potential therapeutic benefit."[83, 84] If a target cannot be validated, it will not proceed in the drug development process.

ii. **Hit and lead generation and optimization** identify compounds or other treatment types with a desired biological activity that could produce an intended therapeutic response in conjunction with a validated target.[85] This is followed by refinement to maintain favorable properties in lead compounds while improving on structural deficiencies. The goal of this step is to identify a compound for pre-clinical testing.

b. For medical devices[86] (e.g., diagnostics, some behavioral interventions as described below):

i. **Design and engineering** are the process of creating a concept or idea for a new device.[87] From here, researchers identify the steps needed to determine whether the concept is workable. The concept can then be built upon and refined through prototypes.
Note: Some software-based behavioral interventions are medical devices under FDA's statute, whereas others, such as those software functions that are "intended for maintaining or encouraging a healthy lifestyle" and are "unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition," are not. See sections 201(h) and 520(o)(1)(B) of the FD&C Act. Please see the Healthcare Delivery chapter for more information on research into non-device behavioral interventions.

3. **Pre-clinical testing** refers to *in vitro* and *in vivo* studies and is designed to advance potential therapeutics for human clinical research further.[88] This is often done to determine any toxic or adverse effects before trials can be carried out in humans and ultimately be made available on the market.[89] If a drug or device shows potential benefits, an investigator can submit to the FDA an investigational new drug application (drugs) or an investigational device exemption application (devices) to proceed to clinical trials, which are discussed in more detail in the Medical Product Development, Safety, and Effectiveness chapter.[90, 91]

AI Risks in Medical Research and Discovery

Because medical research and discovery comprise precursor steps to the use of products and care delivery, any bias or other unaccounted-for risks from AI models leveraged in these steps could be propagated downstream, potentially reaching patients. It is, therefore, critical to consider, manage, and ultimately mitigate associated AI risks. Furthermore, it may be difficult to see adoption at scale without developing trustworthiness in the eyes of patients, caregivers, and providers concerning AI in research and technology. Engaging these communities proactively as the technology develops rapidly could be essential to fostering the safe adoption of these technologies. While the potential is large, future success will depend on how key actors work together to balance risk and manage uncertainty.

Before detailing additional AI benefits and risks in medical research and discovery later in the chapter, three focus areas for managing risks are highlighted: biosecurity, data security, and AI hijacking. It is important to note that these risks are not yet fully understood and may evolve as technology advances, making it difficult to stratify and prioritize them against other risks.

---

82 https://www.fda.gov/media/167973/download
83 https://www.ncbi.nlm.nih.gov/books/NBK195048/
84 https://www.ncbi.nlm.nih.gov/books/NBK195039/
85 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3058157/
86 See Appendix A: "Glossary of terms" for the definition of "medical device" used in this Plan.
87 https://www.fda.gov/patients/device-development-process/step-1-device-discovery-and-concept
88 https://www.fda.gov/media/167973/download
89 https://toolkit.ncats.nih.gov/glossary/preclinical-studies/
90 https://www.fda.gov/drugs/types-applications/investigational-new-drug-ind-application
91 https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/investigational-device-exemption-ide

1. **Biosecurity risks:** In May 2024, the Executive Office of the President released the U.S. Government Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential,[92] which articulates potential applications for the dual use of AI. This includes research conducted for legitimate purposes that generate knowledge, information, technologies, and products that can be utilized to improve care outcomes or research conducted for malicious purposes that could generate potentially harmful bioweapons or harmful pathogens, which present a biosecurity threat to the U.S. and the world. Action has already been taken to help mitigate this threat (see details in section Action Plan), and going forward, HHS and the U.S. government security apparatus can continue to coordinate closely with the research community, private companies (including manufacturers), and the publishing industry to build on the existing guidance from the Executive Office of the President and continue to work to strike the right balance between open science and public security.[93]
2. **Data security risks:** The October 2024 White House Memorandum on Advancing the United States's Leadership in Artificial Intelligence[94] noted some particular risks in medical research in discovery: AI systems leveraged in the process may reveal aspects of their training data—either inadvertently or through deliberate manipulation by malicious actors—causing data spillage from models that may be trained on classified or controlled information when used on networks where such information is not permitted. Going forward, HHS will explore what policy and technical support are needed to ensure the responsible and safe use of these data in AI research and development.
3. **AI hijacking:** Malicious actors can hijack AI models and systems in medical research and discovery contexts by seizing control of agents or solutions to direct them toward harmful actions.[95] This might be particularly relevant to AI use cases in basic research that analyzes large biomedical datasets or in the design and manipulation of drugs or devices. AI hijacking can include poisoning training data.[96] Because AI hijacking can result in breaches of personal health information, controlled or confidential information, and proprietary or national security information, it is a cross-cutting risk and therefore is not listed across each use case in the following table.

1.5.1 Example Use Cases and Risks of AI across the Medical Research and Discovery Value Chain

In the tables below, HHS highlights a non-exhaustive list of potential benefits and risks[97] of AI across the medical research and discovery value chain. Please note that the use cases detailed below highlight existing or potential ways that AI can be used by a variety of stakeholders in this domain. For details on how HHS and its divisions are using AI, please reference the HHS AI Use Case Inventory 2024.[98]

[92] https://aspr.hhs.gov/S3/Documents/USG-Policy-for-Oversight-of-DURC-and-PEPP-May2024-508.pdf

[93] https://aspr.hhs.gov/S3/Pages/OSTP-Framework-for-Nucleic-Acid-Synthesis-Screening.aspx

[94] https://www.whitehouse.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/

[95] https://ieeexplore.ieee.org/document/9131724

[96] https://pmc.ncbi.nlm.nih.gov/articles/PMC10984073/

[97] https://osp.od.nih.gov/policies/artificial-intelligence/

[98] https://www.healthit.gov/hhs-ai-usecases

Functional component 1: Basic research

Advances general knowledge or understanding of biological phenomena and observable facts

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Advanced generative and analytical models that can accelerate the timeline to breakthrough discoveries and expand inventories of potential hypotheses *E.g., analyzing medical texts and other data sources to generate novel biological insights* Analysis and synthesis of significant amounts of information from existing scientific research, publications, and other data sources leveraging AI[99] *E.g., analysis of repositories of large biological datasets to create and refine hypotheses to explore* Advanced processing of large datasets to better understand a condition, biological mechanism, or other health topic can increase the likelihood of a breakthrough discovery[100] *E.g., analysis of potential disease genes, RNA, and proteins involved in disease* Foundational models that can analyze large volumes of genetics data and use ML to identify which biomolecules might be involved in disease[101]	**Bias and validity—potential to introduce bias or produce inaccurate results** *E.g., insights that are not generalizable due to analyzing biased or low-quality data* The results of AI-driven basic research may only be as good as the analyzed data. If datasets do not sufficiently represent the population, results may not be generalizable. This bias can then be propagated throughout the rest of the medical research and discovery value chain, even making its way into medical products used in clinical trials and more. Additionally, there can be potential nefarious manipulation of data or model quality through data poisoning, in which an attacker alters training data to cause AI to "behave in an undesirable way," which could impact the validity and accuracy of results.[102] *E.g., hypotheses that do not accurately reflect data or literature* Poor data quality, management, and/or oversight from investigators not necessarily well-versed in AI could lead to insight generation that is not reflective of reality.
	Privacy, safety, and transparency—potential confidential, sensitive, classified, or personal data breaches or unauthorized disclosures *E.g., intentional or unintentional release or re-identification of personal or confidential information* AI models can potentially be trained on confidential or other sensitive data that may create risks of leaking information that would otherwise be kept private. As a specific example, if training data contains clinical images and/or medical records that are protected health information [PHI]),[103] data breeches can result in PHI being used for training made available to AI users, leading to potential regulatory and policy concerns (e.g., HIPAA).[104] Additionally, as the amount of data collected and analyzed by models increases, even if data is originally de-identified, so does the risk of bad actors (intentionally) or even algorithms (unintentionally) re-identifying knowing or unknowing participants. When integrating multiple datasets or models, data that was otherwise de-identified in each, when combined, may be re-identifiable. Furthermore, consent issues can arise when an AI model uses PHI in one analysis, for which authorization was obtained from patients, is accidentally or intentionally used in subsequent AI analyses not authorized by patients. Such a risk may require new consent and authorization frameworks and more transparency in the future. *E.g., lack of transparency on how clinical data, which may include personal data, could be used in basic research* AI models leveraged in or to inform basic research could use identifiable or de-identified patient data (e.g., to train disease models). The people whose data could be leveraged may not know how their data is used or disclosed, the corresponding potential impacts of that use and disclosure, and any accompanying risks. Mechanisms for appropriate authorization and transparency regarding data use will become increasingly important with increasing AI adoption in basic research.

[99] https://www.fda.gov/media/167973/download

[100] https://pmc.ncbi.nlm.nih.gov/articles/PMC9501106/

[101] https://scopeblog.stanford.edu/2022/06/10/using-ai-to-find-disease-causing-genes/

[102] https://pmc.ncbi.nlm.nih.gov/articles/PMC10984073/

[103] See Appendix A: "Glossary of terms" for the definition of "protected health information (PHI)" used in this Plan.

[104] https://www.hhs.gov/hipaa/for-professionals/index.html

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
	E.g., model card inaccuracy as datasets and models are integrated One approach to AI transparency is to leverage model cards that describe model quality (e.g., data trained concerning demographics, time, quantity, and geography).[105] As models and/or their associated datasets become integrated, their corresponding model cards may lose their accuracy because linked data and models can increase risks related to privacy, re-identifying information, and more. Note that this risk may apply to multiple parts of the value chain but is described here due to the large datasets associated with AI use cases in basic research.

Functional component 2: Discovery

Scientific exploration to find therapies or develop products that may treat or cure disease, which can vary by type of medical product.

See the above discovery description for more details on the type of medical product.

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive models that can leverage basic research insights to predict and prioritize potential therapeutic targets and leads *E.g., analysis of systems biology to predict targets* Using advanced analytics on structural and systems biology knowledge and available genomic, transcriptomic, proteomic, and other data sources from healthy persons and those with a specific disease of interest[106] to predict novel targets[107] *E.g., analysis of drug-target interactions to help facilitate discovery through drug repurposing* Exploration of drug-target interactions that help provide predictions about classes of drugs potentially interacting with the same targets or having a similar mechanism of action, which may help predict the toxicity of a molecule based on specific known features. This strategy can help guide drug repurposing efforts that could utilize previously characterized compounds. Drug repurposing efforts utilizing AI can also potentially benefit from the increased availability of suitable RWD from various sources (e.g., electronic health records (EHRs), registries, and DHTs) to identify previously unknown effects of drugs on disease pathways.[108]	**Bias and validity—potential to introduce bias or produce inaccurate results** *E.g., target identification lead generation based on non-representative datasets and covert AI social bias* Models trained on poor-quality or non-representative datasets (e.g., biomarkers and biomolecules sourced from unbalanced racial or gender demographics) can lead to the identification of targets and leads that apply to only some populations, potentially perpetuating social bias and exacerbating health inequities and group harms. While models have learned how to improve upon biases built through the data they are trained on, research has shown that covert biases are just as, if not more, present, which can exacerbate health inequities and be more difficult to track.[111] *E.g., statistical and computational bias stemming from heterogenous or incorrect data* In AI systems, statistical and computational bias can be present in the datasets and algorithmic processes used to develop AI applications. It can arise when algorithms are trained on one data type and cannot extrapolate beyond that data. The error may be due to heterogeneous data, representation of complex data in simpler mathematical representations, wrong data, algorithmic biases such as over- and under-fitting, the treatment of outliers, and data cleaning and imputation factors.[112] *E.g., inaccurate identification of compounds or devices* Content generated by some AI (e.g., LLMs) can, by design, be based on information directly or inferred indirectly (often referred to as

[105] https://pmc.ncbi.nlm.nih.gov/articles/PMC9284683/

[106] https://www.fda.gov/media/167973/download

[107] https://pmc.ncbi.nlm.nih.gov/articles/PMC7591760/

[108] https://www.fda.gov/media/167973/download

[111] https://hai.stanford.edu/news/covert-racism-ai-how-language-models-are-reinforcing-outdated-stereotypes

[112] https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
E.g., recommendation of research targets and leads *In silico* drug design that enables researchers to predict antibody structures rapidly, assess the structure and function of amino acid mutagenesis, and accelerate *de novo* protein design (e.g., validating oncology targets via GenAI)[109] *E.g., design of nucleic acid and amino acid sequences with specific desired functions* Leveraging AI platforms to create biomolecules with helpful functionality can increase efficacy and speed of drug development[110]	"hallucination"), which introduces potential for inaccuracies that are presented as accurate, sometimes even generating further inaccurate information that justifies inaccuracies when probed to explain further. Using AI that does not aim to reduce this phenomenon algorithmically (e.g., through retrieval-augmented generative models) could introduce this risk to medical research and discovery pipelines and propagate inaccuracies throughout the value chain if not otherwise appropriately solved for (e.g., with a human in the loop). *E.g., unnecessary depletion of resources directed at unfounded targets or leads* Hallucinations or other inaccuracies in AI analyses and predictions related to target identification or hit and lead generation and optimization can deplete financial and/or computational resources on targets or leads that are potentially unsuitable for further exploration.
***In silico* experimentation technologies that can predict behavior, design and manipulate products, and screen drug candidates for effectiveness** *E.g., protein folding prediction to aid in the design of products* Models that can predict the structure of proteins based on large repositories of data using deep learning[113] *E.g., design and manipulation of biomolecules and medical devices* *In silico* experimentation on the structure of biomolecules (e.g., DNA, RNA, and proteins) for testing candidate drugs and MoAs or on the structure of medical devices to help determine potential applicability before pre-clinical studies[114] *E.g., drug compound screening* Prediction of the chemical properties and bioactivity of compounds and their efficacy and potential adverse events based on the compound's specificity and affinity for a target[115]	**Biosecurity threats—potential to create harmful products** *E.g., malicious or unintentional design of novel pathogenic or toxic biological and chemical agents, including nucleic acid sequences, proteins, and peptides* Using AI on publicly available research data or leveraging design and folding AI technologies could be conducted for legitimate or malicious purposes and may generate—more easily than through traditional research activities that don't use AI—novel pathogenic or toxic agents that are not currently addressed by research oversight frameworks, such as the 2024 U.S. Government Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential.[116] DNA and RNA sequences of these agents may also not be detected by the current best match criteria in the OSTP Framework for Nucleic Acid Synthesis Screening, and others (e.g., proteins, peptides) may be able to defeat natural immune systems or existing medical interventions to treat disease.[117]

[109] https://pubmed.ncbi.nlm.nih.gov/35679624/

[110] https://www.ucsf.edu/news/2023/01/424641/ai-technology-generates-original-proteins-scratch

[113] https://directorsblog.nih.gov/2021/07/27/artificial-intelligence-accurately-predicts-protein-folding/

[114] https://www.nature.com/articles/s41392-023-01381-z

[115] https://www.fda.gov/media/167973/download

[116] https://www.whitehouse.gov/wp-content/uploads/2024/05/USG-Policy-for-Oversight-of-DURC-and-PEPP.pdf

[117] https://www.whitehouse.gov/wp-content/uploads/2024/04/Nucleic-Acid_Synthesis_Screening_Framework.pdf

Functional component 3: Pre-clinical testing

Investigations that evaluate a drug, procedure, or medical device in cell and/or animal models to determine any toxic or adverse effects before trials can be conducted in humans.

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive models, analytical devices, and representation tools that accelerate timelines to care and bolster understanding of discoveries before going to trial *E.g., prediction of drug and device efficacy and safety to determine fit for trials* Multimodal data-based (e.g., registries, omics, knowledge graphs, RWD) comparisons of potential efficacy before clinical trials to mitigate risk and potentially save significant trial costs for potential failures[118] *E.g., medical imaging analysis of in vivo and in vitro testing* Automated analysis of research images for identifying structures to help select drugs for clinical trials[119] *E.g., digital twins to increase diversity and sample size of in vivo and in vitro tests* Virtual representations of objects, systems, or animal candidates can accelerate and strengthen pre-clinical research by enabling additional simulated testing[120] *E.g., life sciences workflow optimizations* ML can be used to "predict millions of workflow configurations" and optimize them to run as efficiently as possible on distributed computing data infrastructure, enabling faster discovery.[121]	**Validity, bias, and effectiveness, including potential false positives and false negatives** *E.g., unintentionally propagating ineffective ideas or discarding promising solutions* Without a human in the loop to assess the validity of millions (or more) of analyses of identified potential drugs, devices, and research subjects, errors in synthesis and prioritization of outcomes can lead to false positives and negatives in recommended results. *E.g., degradation of model integrity and diverse representation as synthetic data is iterated on* Using synthetic data, even with a positive intent to increase diversity, can erode model quality as it is analyzed and re-analyzed to produce additional synthetic data, and so on. This could jeopardize the accuracy and validity of results and ultimately not achieve the potential goals of increasing representation and/or reducing bias. **Deskilling researchers and investigators**[122] *E.g., reduction of human-led laboratory processes* Automated generation of reliable, safe, and secure laboratory procedures and operations may lower the skill and training requirements for working with high-consequence biological materials, which could lead to the loss of important, highly skilled human talent.
LLMs that can enhance the quality and speed process of regulatory submissions *E.g., generative and analytical regulatory package writing* Using GenAI to develop application materials based on pre-clinical research outcomes for investigational new drug applications and other pre-clinical trial steps that require extensive writing *E.g., digital assistants to automate procedures and analyses* Agent assistants that maintain, analyze, and synthesize outputs from scientific records during experimentation (e.g., ambient listening, the AI Scientist)[123, 124]	**Potential lack of explainability of research results** *E.g., regulatory submission materials that do not correctly represent outcomes of pre-clinical research* Traceability to the root data used by a model is not always available in AI technologies, which can reduce the verifiability of the results or intermediate conclusions of its outputs. This potential lack of validity can reduce stakeholders' trust in results (e.g., academia, industry, the general public, and regulators).

[118] https://pmc.ncbi.nlm.nih.gov/articles/PMC10720846/
[119] https://pmc.ncbi.nlm.nih.gov/articles/PMC7594889/
[120] https://pubmed.ncbi.nlm.nih.gov/37030076/
[121] https://www.anl.gov/article/accelerating-discovery-optimizing-workflows-to-advance-the-use-of-ai-for-science
[122] Note that a conceptually similar risk in the context of AI use by clinicians is discussed in the Healthcare Delivery chapter.
[123] https://www.nature.com/articles/d41586-024-02842-3
[124] https://pubmed.ncbi.nlm.nih.gov/35584760/

There are opportunities to develop and employ AI to improve medical research and discovery quality, quantity, and speed. From AI that supports focusing on hypotheses through target identification and optimization at the lab bench to analyzing large datasets, there is strong evidence for optimism. However, this enthusiasm must be balanced by the reality that these applications have risks that deserve careful attention and mitigation strategies. Every stakeholder must monitor and mitigate risks. HHS will use the following action plan to empower entities and individuals across the value chain to increase their adoption of AI safely, responsibly, equitably, and impactfully.

1.6 Action Plan

In light of the evolving AI landscape in medical research and discovery, HHS has taken multiple steps to promote responsible AI use by providing resourcing to intramural and extramural research, advancing accessibility of streamlined datasets, developing workforce talent and capabilities, and many other actions to date. The Action Plan below follows the four goals that support HHS's AI strategy: 1. catalyzing health AI innovation and adoption; 2. promoting trustworthy AI development and ethical and responsible use; 3. democratizing AI technologies and resources; and 4. cultivating AI-empowered workforces and organization cultures. For each goal, the Action Plan provides context, an overview of HHS and relevant other federal actions to date, and specific near- and long-term priorities HHS will take. HHS recognizes that this Action Plan will require revisions over time as technologies evolve and is committed to providing structure and flexibility to ensure longstanding impact.

1.6.1 Catalyze Health AI Innovation and Adoption

Increasing AI adoption in medical research and discovery can transform the quality and speed of innovation that ultimately improves patient outcomes. HHS has an opportunity to increase AI adoption by pursuing the following themes of actions:

1. Expanding the breadth of medical research and discovery AI use across disease areas and steps of the value chain
2. Enhancing coordination across geographies to harness AI to improve medical research and discovery
3. Fostering AI-ready data standards and datasets to bolster their usability for AI-empowered medical research and discovery

Below, HHS discusses the context of each theme of action in more detail, corresponding actions to date, and plans to promote AI innovation and adoption in medical research and discovery.

1. **Expanding the breadth of medical research and discovery AI use across disease areas and steps of the value chain:**

   Context:

   AI's relatively higher uptake in discovery (e.g., *in silico* target identification, high-throughput screening of potential candidates) than in other parts of the value chain, coupled with the potential incentives AI faces to focus on disease areas with higher market potential, indicates an opportunity to catalyze further AI adoption by focusing on AI use cases across other parts of the value chain (e.g., basic research, preclinical studies) and in the exploration of more disease areas (e.g., less researched, those with high unmet needs). HHS is focused on expanding applications of AI in medical research and discovery while maintaining integrity in its resourcing programs, which can include resourcing, training, or additional policies or guidelines. HHS will

look to advance AI adoption that could help meet unmet patient needs and foster innovation across the full value chain more broadly.

HHS actions to date (non-exhaustive):

- **National Cancer Institute's (NCI)**[125] **Informatics Technology for Cancer Research** funds research-driven informatics technology across the development life cycle to address priority needs in cancer research.[126] These projects are increasingly developing or incorporating advanced AI methods. The program supports the development of critical tools and resources to improve the acquisition, analysis, visualization, and interpretation of data across the cancer research continuum, including cancer biology, cancer treatment and diagnosis, early cancer detection, risk assessment and prevention, cancer control and epidemiology, and cancer health equity.
- **National Institute of Mental Health's (NIMH's) Theoretical and Computational Neuroscience Program** supports basic experimental and theoretical research focusing on biophysically realistic computational approaches modeling dynamical processes in the brain, from single cell activity to neural systems regulating complex behaviors.[127]
- **NIMH's Translational Digital and Computational Psychiatry Program** fosters innovative computational approaches to identify and validate novel mechanisms, biomarkers, and treatment targets for preventing and treating psychiatric disorders. The program supports research projects that use advanced computational methods with behavioral, biological, and/or clinical data to decipher complex mechanisms involved in mental disorders and to conduct initial tests of novel tools to predict risk, clinical trajectories, and treatment response.[128]
- **The ARPA-H TARGET program** will expand the pool of candidate molecules with antibiotic potential using deep learning to filter for candidate biomolecules and GenAI to broaden the scope of possible pharmaceuticals.[129]
- **ARPA-H's Computational ADME-Tox and Physiology Analysis for Safer Therapeutics (CATALYST) program** envisions a future where approval to begin first-in-human clinical trials can be based on *in silico* safety data.[130] The program focuses on developing animal-free, sound experimental practice methods with specific attention to pharmacokinetics, including absorption, distribution, metabolism, and excretion (ADME), and pharmacodynamics for safety and toxicity. CATALYST will pursue novel technologies that reliably represent human physiology to reduce the failure rate of investigational new drug candidates. Such technologies will ensure that medicines reaching clinical trials have confident safety profiles and better protect diverse trial participants and future patients.
- **NIH's Brain Research Through Advancing Innovative Neurotechnologies® (BRAIN) Initiative: Theories, Models, and Methods for Analysis of Complex Data from the Brain** develops theories, computational models, and analytical tools to derive the understanding of brain function from complex neuroscience data. Proposed projects could develop tools to integrate existing theories or formulate new theories; conceptual frameworks to organize or fuse data to infer general principles of brain function; multiscale/multiphysics models to generate new testable hypotheses to design/drive future experiments; new analytical methods to substantiate falsifiable hypotheses about brain function. The tools developed were expected to be widely available for use and modification in the neuroscience research community.[131]

---

[125] Note that NCI is a subsidiary of NIH.

[126] https://www.cancer.gov/about-nci/organization/cssi/research/itcr

[127] https://www.nimh.nih.gov/about/organization/dnbbs/behavioral-science-and-integrative-neuroscience-research-branch/theoretical-and-computational-neuroscience-program

[128] https://www.nimh.nih.gov/about/organization/dtr/adult-psychopathology-and-psychosocial-interventions-research-branch/translational-digital-and-computational-psychiatry-program

[129] https://arpa-h.gov/news-and-events/arpa-h-project-accelerate-discovery-and-development-new-antibiotics-using

[130] https://arpa-h.gov/research-and-funding/programs/catalyst

[131] https://grants.nih.gov/grants/guide/rfa-files/RFA-DA-23-039.html

HHS near-term priorities:

- Explore resourcing for medical research and discovery leveraging AI to address TAs with unmet needs and/or identify and analyze novel rather than known targets.
- Explore resourcing research, training, and workshops focusing on basic and pre-clinical research areas with lower AI adoption, such as late-stage investigations closer to the regulatory approval process.
- Continue to hold webinars, workshops, listening sessions, and more to socialize notices of funding opportunities (NOFOs) and requests for information.[132]
- Identify barriers to the adoption of AI across the value chain.
    - Convene stakeholders to delineate technical, economic, workforce, data availability, and regulatory hurdles to adopting AI across the medical research and discovery value chain.
    - Convene patients and other stakeholders to address transparency and build trust (see "enabling risk management and transparency of AI" under "Promote Trustworthy AI Development and Ethical and Responsible Use").
- Explore potential mechanisms to reduce barriers to adoption (e.g., environmental considerations and costs associated with adoption).
- Prioritize and explore resourcing for evidence-building to evaluate responsible AI medical research and discovery investments and maximize the efficacy of HHS spending.
- Provide policy clarity and/or guidelines on acceptable uses of AI in federally funded pre-clinical research (e.g., uses of AI to replace animal-based studies).
- Provide policy clarity and/or guidelines on the uses of AI toward drafting research grant applications and submissions to ensure fairness and transparency and to protect program integrity.
- Adopt AI within HHS to streamline grant review, approval, and support process, subject to robust safeguards to protect program integrity, equity, and fairness.

HHS long-term priorities:

- Explore experimentation opportunities regarding economic frameworks for exchanging data and AI models that can make pricing affordable while allowing for fair compensation and safety of AI use.

2. Enhancing coordination across geographies to harness AI to improve medical research and discovery:

Context:

Multiple bodies internationally and in the U.S. have varying regulations that could impact the medical research and discovery space (e.g., General Data Protection Regulation, European Union AI Act, and HIPAA).[133] Coordination between these bodies on their approach to AI in the context of medical research and discovery could reduce barriers to innovation while still maintaining the safety and efficacy of corresponding use cases. Without proactive coordination, achieving realizable improvements in these areas will be diffuse and suboptimal given the complexity of the value chain and underlying economics and the considerable number of public and private sector stakeholders involved. HHS can bolster future innovation by engaging stakeholders—domestically and abroad—to promote further alignment across the value chain.

[132] All materials must be digitally accessible and webinars and listening sessions must, at a minimum, have ASL interpreters. If recorded, the recording needs closed captions and audio descriptions. Furthermore, the NOFO and RFIs must include digital accessibility language to ensure all materials provided are conformant.

[133] https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/

HHS actions to date (non-exhaustive):

- **The NIH Common Fund's Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program**[134] leverages data science technologies and prior NIH investments to develop solutions to the continent's most pressing public health problems through a robust ecosystem of new partners from academic, government, and private sectors.

HHS near-term priorities:

- Prioritize and explore resources for the most promising collaborative, multidisciplinary, and cross-border proposals for AI integration in basic and pre-clinical research.
- Facilitate coordination across HHS divisions to share appropriate data, methodology, technologies, and resources related to medical research and discovery to enable stronger HHS innovation activities.

HHS long-term priorities:

- Define and establish policies and guidelines for cross-border AI in medical research and discovery collaboration that comply with U.S. standards.
- Provide guidelines to other agencies and STLTs related to AI and data-sharing standards, as appropriate and authorized within HHS domains,[135] to enhance the possibility of stronger international collaboration in medical research and discovery.

3. Fostering AI-ready data standards and datasets to bolster their usability for AI-empowered medical research and discovery:[136]

(See Goal 3: "Democratize AI Technologies and Resources" theme of action 2: "Increasing accessibility to responsibly curated AI-ready data tooling and infrastructure for those who are less able to access them today" for more information on data infrastructure and tooling)

Context:

Variability in the quality, volume, and representativeness of data used for training AI could lead to its underperformance due to bias and shortcut learning.[137] While the healthcare delivery system generates a tremendous amount of clinical and administrative data, fragmentation of the industry poses considerable challenges to the aggregation of high-quality data for AI model development to support pre-clinical medical research and discovery. Additionally, models trained on clinical data that contain personal information are difficult to share broadly. Furthermore, proprietary or confidential molecular, chemical, and other non-clinical data could be fragmented across industry and academia. As a result, vast amounts of data that could be used for research cannot be easily tapped. By focusing on making this data AI-ready for medical research and discovery, HHS can empower further AI adoption in the space.

HHS actions to date (non-exhaustive):

- **NIH's Bridge2AI program** funds studies to generate flagship datasets and best practices for the collection and preparation of AI-ready data to address biomedical and behavioral research challenges (e.g., generating new flagship biomedical and behavioral datasets that are ethically sourced, trustworthy, well-defined, and accessible, developing software and standards to unify data attributes across multiple

[134] https://commonfund.nih.gov/AfricaData

[135] https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-119-1.pdf, https://www.govinfo.gov/app/details/PLAW-104publ113. Under OMB Circular A-119 and the National Technology Transfer and Advancement Act of 1995 (Public Law 104-113), NIST has primary authority to coordinate standards, with reservations for other Federal functions with specific authority for domain-specific standards. That said, HHS agencies do have domain-specific standards.

[136] This aligns with the 2024-2030 Federal Health IT Strategic Plan Goal 2: Enhance the Delivery and Experience of Care, Objective D: Providers experience reduced regulatory and administrative burden, Strategy: Promote the safe, secure, and responsible use of AI tools and standards so that healthcare providers and patients can expect trustworthy, relevant, and representative results from AI tools that provide better, more streamlined care delivery.

[137] https://www.nature.com/articles/s41746-024-01118-4 "Shortcut learning refers to a phenomenon in which an AI model learns to solve a task based on spurious correlations present in the data as opposed to features directly related to the task itself."

data sources and data types, creating automated tools to accelerate the creation of FAIR [Findable, Accessible, Interoperable, and Reusable] and ethically sourced datasets, providing resources to disseminate data, ethical principles, tools, and best practices, creating training materials and activities for workforce development that bridges the AI, biomedical, and behavioral research communities).[138]

- NIH developed **ScHARe,** a cloud-based data platform comprising federated social determinants of health (SDOH) datasets to accelerate research in health disparities, healthcare delivery, health outcomes, and AI bias mitigation strategies.[139]
- NIH, as a part of the **National AI Research Resource (NAIRR) Pilot**[140] with NSF, National Center for Science and Engineering Statistics (NCSES), and the Department of Energy (DOE), leverages large RWD sets to (1) build a synthetic data generator toolkit and framework to assess privacy risk and utility for using such data for evidence-building, and (2) linked medical imaging data with clinical records that will build capacity for multimodal AI development.
- **NIH's BRAIN Initiative: Data Archives** advances research by creating a data archive with appropriate standards and summary information that is broadly available and accessible to the research community for further research. Teams work with the research community to incorporate software tools that allow users to analyze and visualize data and use appropriate standards to describe the data.[141]
- **NIH's BRAIN Initiative: Integration and Analysis of BRAIN Initiative Data** developed informatics tools for analyzing, visualizing, and integrating data related to the BRAIN Initiative or to enhance our understanding of the brain. The tools were user-friendly in accessing and analyzing data from appropriate data archives and could analyze/visualize data without requiring users to download data.[142]

HHS near-term priorities:

- Define and prioritize standards that maximize the findability, accessibility, interoperability, and reusability of research data (including common data elements, metadata, persistent identifiers, and security) with U.S. government partners (e.g., NIST due to their 2024 Research Data Framework [RDaF],[143] United States Core Data for Interoperability [USCDI]) to streamline training and refinement of algorithms with biomedical research data.
- Accelerate alignment of federally funded research data standards (semantic, format, transport) with HHS-adopted standards for EHRs, healthcare providers, and payers (e.g., USCDI,[144] USCDI+,[145] HL7 Fast Healthcare Interoperability Resources [FHIR],[146] CARIN[147]).
- Develop open-source, open-standard tooling and infrastructure for AI data management, cross-standard data mapping, de-identification, etc., to develop AI-ready datasets and tooling.
- Accelerate work with standards development organizations and industry collaborations on standards to support AI development and use across the life cycle.
- Convene a public-private community of practice for sharing best practices regarding data appropriate for AI model use in medical research and discovery, where stakeholders can also collaborate to identify enablers/barriers to access such data.
- Explore potential safe ways to leverage and share AI models trained on clinical or other personal information without risking privacy, consent, or transparency.
- Accelerate federated ML research, tooling, and implementation support; facilitate a public-private process to define open-industry standards and conventions for federated ML.

---

138 https://commonfund.nih.gov/bridge2ai

139 https://www.nimhd.nih.gov/resources/schare/

140 https://nairrpilot.org/

141 https://grants.nih.gov/grants/guide/rfa-files/RFA-MH-25-110.html

142 https://grants.nih.gov/grants/guide/rfa-files/RFA-MH-23-270.html

143 https://www.nist.gov/publications/nist-research-data-framework-rdaf-version-20

144 https://www.healthit.gov/isp/united-states-core-data-interoperability-uscdi

145 https://www.healthit.gov/topic/interoperability/uscdi-plus

146 https://www.healthit.gov/sites/default/files/page/2021-04/What%20Is%20FHIR%20Fact%20Sheet.pdf

147 https://www.carinalliance.com/

- Accelerate the development of a research exchange purpose in the **Trusted Exchange Framework and Common Agreement™ (TEFCA™)**[148] to support high-scale, network-facilitated data exchange for research.

HHS long-term priorities:

- Establish the governance, legal, and analytical frameworks as a public resource for AI-ready medical research and discovery datasets.

1.6.2 Promote Trustworthy AI Development and Ethical and Responsible Use

As AI adoption in medical research and discovery continues to advance rapidly, its associated risks may require close attention from HHS to ensure uptake is safe, responsible, and impactful for patients around the world. Key themes of action that HHS could address to ensure the trustworthy and safe use of AI in medical research and discovery include:

1. Building and disseminating evidence to mitigate biosecurity, data security, privacy, and data collection risks
2. Setting clear guidelines for safe and trustworthy AI use in medical research and discovery and the distribution and use of federal resources
3. Enabling safe and responsible organizational governance of AI risk management and transparency

Below, HHS discusses the context of each theme of action in more detail, corresponding actions to date, and plans to ensure the trustworthy and safe use of AI in medical research and discovery.

1. **Building and disseminating evidence to mitigate biosecurity, data security, privacy, and data collection risks**

   Context:

   As discussed in Section 1.5.1, AI in medical research and discovery could be used nefariously to create biosecurity and biosafety threats (e.g., potential novel pathogens). Additionally, confidential, sensitive, or classified information could be leaked—intentionally or unintentionally—through AI model training and deployment, and collecting sensitive patient data could require de-identification or authorization from patients, both of which can present challenges to gathering statistically powerful quantities of information for medical research and discovery.

   The HIPAA Privacy Rule has specific provisions related to the use and disclosure of patient information for research[149] (Note that the HIPAA Privacy Rule has provisions related to use and disclosures of PHI for a variety of circumstances which are further outlined in the Healthcare Delivery chapter), and AI models present unique considerations regarding adherence with privacy protections. Potential patient concerns include lack of consent for the use of their de-identified data and transparency into how their consented personal data are used. AI makes it easier to re-identify information leveraging various datasets, including publicly available external data, which may require the adjustment of data-sharing policies and practices, especially with entities not subject to HIPAA. HHS and the federal government have taken action to approach this, and going forward, HHS will pursue further actions to continue protecting sensitive information regarding AI use in medical research and discovery.

[148] https://www.healthit.gov/topic/interoperability/policy/trusted-exchange-framework-and-common-agreement-tefca

[149] https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html

HHS actions to date (non-exhaustive):

- **NIH's Data Management and Sharing Policy** promotes the sharing of scientific data to help accelerate biomedical research discovery, in part, by enabling validation of research results, providing accessibility to high-value datasets, and promoting data reuse for future research studies. It also emphasizes the importance of good data management practices and establishes the expectation for maximizing the appropriate sharing of scientific data generated from NIH-funded or conducted research, with justified limitations or exceptions. [150]
    - **NIH's Data Management and Sharing Policy Supplemental Information on Protecting Participant Privacy When Sharing Human Scientific Data** outlines principles, best practices, and points to consider for researchers to protect the privacy of research participants when sharing participant data. The framework does not establish binding rules but rather provides a framework for sharing both identifiable and de-identified data as well as data obtained with consent and data where consent was not required.[151]
- **Implementation of the Executive Office of the President's National Biodefense Strategy,**[152] which explains how the U.S. Government will manage its activities more effectively to assess, prevent, protect against, respond to, and recover from biological threats, which could implicitly incorporate threats from AI use.
- **HHS's Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids** describes its screening framework guidance, which sets forth baseline standards for the gene and genome synthesis industry, as well as best practices for all entities involved in the provision, use, and transfer of synthetic nucleic acids regarding screening orders and recipients and maintaining records.[153] In addition, this guidance seeks to encourage best practices to address biosecurity concerns associated with the potential misuse of synthetic nucleic acids in order to bypass existing regulatory controls and commit unlawful acts.
- **Implementation of the Executive Office of the President's Framework for Nucleic Acid Synthesis Screening,**[154] which is consistent with and responsive to the guidance in the HHS Screening Framework and fulfills provisions in the 2023 Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence that requires all researchers receiving U.S. government life sciences research funding to procure synthetic genetic materials only from companies that comply with sequence screening best practices (88 FR 7519).[155]
- **HHS's HIPAA Privacy Rule** establishes the conditions under which PHI may be used or disclosed by covered entities for research purposes (45 CFR part 160 and subparts A and E of part 164).[156] Under this Privacy Rule, covered entities are permitted to use and disclose PHI for research with individual authorization or without individual authorization under limited circumstances set forth in the Privacy Rule. While the Privacy Rule may not explicitly discuss AI, its safeguards apply whether AI is leveraged in medical research and discovery or not.
- **The Belmont Report**, written by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, is a statement of basic ethical principles and guidelines that should assist in resolving the ethical problems that surround the conduct of research with human subjects, which can apply regardless of the technologies being used in research and discovery, including but not limited to AI in medical research and discovery analyzing clinical data.[157]

[150] https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

[151] https://sharing.nih.gov/data-management-and-sharing-policy/protecting-participant-privacy-when-sharing-scientific-data

[152] https://aspr.hhs.gov/biodefense/Pages/default.aspx

[153] https://aspr.hhs.gov/legal/synna/Documents/SynNA-Guidance-2023.pdf

[154] https://aspr.hhs.gov/S3/Documents/OSTP-Nucleic-Acid-Synthesis-Screening-Framework-Sep2024.pdf

[155] https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

[156] https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html

[157] https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

HHS near-term priorities:

- Iteratively monitor and evaluate potential nefarious uses to continuously refine guidelines and policies related to biosecurity and data breeches.
- Consider vetting predictive methodologies for use in amino and nucleic acid sequence screening per the **Screening Framework Guidance.**[158]
- Facilitate the public-private process to define open industry standards to accelerate the availability of privacy-enhancing technologies for data de-identification (e.g., privacy-preserving record linkage (PPRL), differential privacy).
- Evaluate potential technical solutions that would allow developers and investigators to create and use models in a sandbox[159] environment that would prevent data spillage to enable the safe testing and progression of AI use in medical research and discovery.
- Explore the opportunities and risks of leveraging AI in data collection, including the quality of the data (e.g., EHRs potentially showcasing high-quality versus low-quality outcomes in some clinical settings versus others).
- Explore potential data use authorization pathways that enable the use of patient data in iterative and potentially multi-use AI models while maintaining protections consistent with HHS values, regulations, and policies.
- Explore resourcing for the evaluation of homomorphic encryption and data security, which enable the federation of data without allowing visibility into data linkages, for the safe use of AI in medical research and discovery settings.
- Explore approaches to protect AI models used in medical research and discovery and sensitive health data from adversarial attacks.
- Explore the development of mechanisms to prevent and reduce harm from the misuse of predictive analytics tools used in medical research and discovery.
- Provide guidelines on training models on patient, participant, genomics, and controlled access data since there is a high risk of data breach and privacy and confidentiality concerns. Consider soliciting community input to inform these guidelines.
- Explore data-sharing protocols that protect sensitive health information.

HHS long-term priorities:

- Consider potential policy solutions or guidelines that enable medical research and discovery to leverage AI outside of controlled access environments while minimizing the risk of data spillage.
- Provide policy clarity and/or guidelines on special considerations regarding AI in research, including definitions of AI developed specifically for research, usability for research of AI models, re-identification risks of patient data used and shared for research, and privacy and security implications for AI in research contexts.
- Evaluate potential pathways to engage STLTs on common pathways for patients to authorize their data use in medical research and discovery to enhance diversity and representation in medical research and discovery while also designing long-term solutions to accelerate and amplify safe data collection and use.
- Consider potential technical or policy solutions that minimize barriers to patient data collection while upholding data security and minimizing unauthorized use.

[158] https://aspr.hhs.gov/legal/synna/Documents/SynNA-Guidance-2023.pdf

[159] See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan.

2. **Setting clear guidelines for safe and trustworthy AI use in medical research and discovery and the distribution and use of federal resources**

Context:

Establishing and fostering trustworthy AI is paramount to the responsible adoption of AI in medical research and discovery. Developing evidence for and disseminating guidelines and regulatory expectations related to transparency and other ethical, legal, and social implications (ELSI) of AI models used in medical research and discovery, including those that leverage federal resources, may lead to safer and more trustworthy use of AI in the space. HHS has taken steps to address this challenge and will continue to build safeguards in the future.

HHS actions to date (non-exhaustive):

- HHS policymakers have established a **regulatory framework, known as the Common Rule, to guide biomedical research.** This framework will continue to support the ethical and responsible use of AI throughout the research life cycle.[160] Appendix B includes specific web pages detailing how these regulations, policies, and best practices should be considered before, during, and after the development and use of AI in research. The **main tenets of this policy framework** include:
  - **Protection of human subject research participants,** which aims to safeguard research participants' rights, safety, and welfare.
  - **Health information privacy** policies, regulations, and best practices help protect the privacy and security of health data used in research, thereby fostering trust in healthcare research activities.
  - **Biosecurity and biosafety oversight** that continues to apply to the development or use of AI in biomedical research.
  - Policy and guidance around **public access to research products and data management and sharing,** which seek to maximize the responsible and appropriate sharing and management of research products while ensuring that researchers consider how human research participants' privacy, rights, and confidentiality will be protected. Responsible and appropriate sharing and management refer not exclusively to human data protections but also to other relevant laws, regulations, and policies that limit disclosure and restrictions on sharing imposed by agreements.
  - **Licensing, intellectual property, and technology transfer policy** and resources related to intellectual property and software sharing to complement data sharing and delineate investigator rights.
- **NIH's Artificial Intelligence in Research Policy Considerations and Guidance** details a robust system of policies and practices that guide stakeholders across the biomedical and behavioral research ecosystem.[161] NIH's policy framework is designed to responsibly guide and govern advancing science and emerging technologies, including developing and using AI technologies in research. The policies, best practices, and regulations discussed reflect this framework and should be considered before, during, and after the development and use of AI in research. It is not an exhaustive list of all policies and requirements that may apply to any NIH-supported research projects. Still, it can guide the research community regarding privacy, intellectual property, data management, participant protection, and more.

HHS near-term priorities:

- Coordinate between midstream (e.g., NIH) and downstream (e.g., FDA) medical research and discovery agencies to enhance information sharing among agencies, where possible, and assist developers aiming to seek regulatory authorization.

[160] https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html

[161] https://osp.od.nih.gov/policies/artificial-intelligence/

- Explore developing a common framework of expectations for addressing or providing transparency into how researchers using AI in medical research and discovery address ELSI in order to proceed to clinical trials and potential regulatory approval.
- Consider supporting guidelines and educational tools to help AI developers as they work toward safety, security, and trust while creating AI technologies for use in medical research and discovery.
- Explore targeting research resources, training, and workshops to further research on the ELSI of AI in medical research and discovery, including explainable AI.
- Create opportunities for communities of practice (e.g., sandboxes)[162] to evaluate ELSI of AI technologies in medical research and discovery internally at a reduced cost.

HHS long-term priorities:

- As necessary, implement updates and/or new policies to ensure responsible use of AI in both internal (e.g., through HHS and/or HHS grant or contract recipients) and external (e.g., in industry and/or academia) medical research and discovery, including potential stratification of AI risks in medical research and discovery.
- Continue prioritizing and exploring resourcing for evidence-building to evaluate ELSI of AI in medical research and discovery as the field continuously evolves.
- Continually monitor advances in AI in medical research and discovery to periodically update and revise policy and/or guidelines to provide further clarity on AI use as it relates to later regulatory approval processes, ELSI, and drug and biological product approval and device marketing authorization requirements.

3. Enabling safe and responsible organizational governance of AI risk management and transparency:

Context:

The trustworthy use of AI relies on the assurance of model performance and characteristics and the implementation and associated workflows that determine how AI is used in practice. There is already considerable policy guidance on responsible research practices covering AI uses.[163] However, a lack of risk management policies targeted specifically to the uses of AI in medical research and discovery may lead to poor AI performance regardless of the quality of the technology.

Additionally, communities can help identify risks pertinent to their residents and align on transparency goals, which could lower the risk of people losing trust in how their data are used.[164] Currently, there are limited standardized approaches for representing the characteristics of AI models used in medical research and discovery to better inform users and regulatory authorities about the potential pitfalls of specific AI models. HHS has approached this challenge by funding and researching such technologies. HHS will continue to share guidelines, develop policy, and explore resourcing activities that support these goals.

HHS actions to date (non-exhaustive):

- **ARPA-H's Performance and Reliability Evaluation for Continuous Modifications and Useability of Artificial Intelligence (**PRECISE-AI**) program** funds investigation to develop technology that can detect when AI used in real-world clinical care settings is out of alignment with underlying training data and, importantly, auto-correct it.[165]

---

[162] See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan.

[163] https://osp.od.nih.gov/policies/artificial-intelligence/

[164] The 2024-2030 Federal Health IT Strategic Plan has a strategy related to this under Goal 1: Promote health and well-being, Objective B: Individuals and populations experience modern and equitable healthcare, Strategy: The federal government plans to promote education, outreach, and transparency about the use of AI technologies and how analysis and outputs of these technologies are applied across the healthcare system so that individuals and healthcare providers are better informed about the use of AI technologies in healthcare, and have transparency into performance, quality, and privacy practices.

[165] https://arpa-h.gov/research-and-funding/programs/precise-ai

- **The Department of Veterans Affairs (VA) and FDA's upcoming collaborative Virtual Health AI Lab** will test medical AI applications in a virtual lab environment to ensure they work, are safe and effective for veterans and patients, and adhere to trustworthy AI principles.[166, 167]
- **HHS's Trustworthy AI Framework** describes what approaches could be taken to address many ethical and other challenges related to AI in healthcare, including those that could apply to medical research and discovery.[168] While not an official policy, it could clarify how HHS approaches addressing these challenges related to AI uptake.
- **AHRQ's Digital Healthcare Equity Framework** guides users in intentionally considering equity in healthcare solutions involving digital technologies and assessing whether these solutions are equitable at every digital healthcare life cycle phase.[169]

HHS near-term priorities:

- Explore the opportunities and risks of leveraging AI in data collection, including the quality of the data (e.g., EHRs showcasing high-quality versus low-quality outcomes).
- Explore synthetic data risk management technical or policy solutions that can reduce the potential degradation of synthetic data as it is iterated on through analyses and subsequent generation of additional synthetic data.
- Develop plans for a quality assurance program for AI used in research aligned with the broader HHS quality assurance policy and program, including digital accessibility for all planning, development, and release.
- Explore strategies to mitigate misuse and approaches to define and assess the risk of current AI models, datasets, and research results.
- In consultation with other federal agencies, update and refine risk management guidelines for federally funded research activities to proactively identify, assess, and mitigate risks associated with AI used in research.
- Define, prioritize, and disseminate frameworks for testing, evaluating, validating, and verifying algorithms used in medical research and discovery.
- Explore opportunities for encouraging transparency of AI model use and personal data use to stakeholders across the value chain whose data may contribute to groundbreaking research, including accompanying risks.
- Train researchers and members of the public who are less skilled, less experienced, and less educated on AI topics to ensure they understand potential dual-use and other risks of AI used in medical research and discovery.[170]
- Explore potential applications of AI to dynamically assess the risk of AI used in medical research and discovery, given the dynamic nature of models and the static current risk management frameworks in place.

HHS long-term priorities:

- Explore privacy-enhancing technologies and their potential use in HHS-supported and HHS-conducted research involving AI.

---

166 https://www.politico.com/newsletters/future-pulse/2024/11/01/a-government-ai-lab-is-born-00186664

167 https://www.nextgov.com/artificial-intelligence/2024/10/va-announces-creation-new-ai-testing-ground-fda/400681/?oref=ng-homepage-river

168 https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf

169 https://digital.ahrq.gov/health-it-tools-and-resources/digital-healthcare-equity/digital-healthcare-equity-framework-and-guide

170 Aligns with 2024-2030 Federal Health IT Strategic Plan Goal 3: Accelerate Research and Innovation, Objective B: Individual and population-level research, analysis, and its application are enhanced by health IT, Strategy: The federal government plans to promote the increased transparency into the development and use of AI algorithms in healthcare settings for providers and patients so that researchers, technology developers, and other health IT users understand how the AI systems work, what kinds of data they are being trained on, and how they are being used in decision-making to mitigate biases, risks, and inaccuracies in AI outputs.

- Partner with industry to develop "research model card" frameworks for standardized representation of characteristics of AI models used in medical research and discovery, including (1) designed purpose, (2) key development inputs, (3) key model outputs, (4) external validation process and results; and (5) life cycle management plan and process.

1.6.3 Democratize AI Technologies and Resources

AI approaches have the potential to "level the playing field" for researchers, helping to identify previously undetectable patterns in extensive, rich, multimodal, and complex datasets, not unlike how CRISPR has made gene editing widely available around the globe. However, access to a broader selection of researchers and applicability to a wider set of underinvested TAs may not happen on their own; federal government direction, incentives, and policies play a key role in ensuring that AI technologies are used for purposes that the market might not adequately or rapidly fulfill on its own (See Goal 1: "Catalyze Health AI Innovation and Adoption" theme of action 1: "Expanding the breadth of medical research and discovery AI use across disease areas and steps of the value chain" for more information). While innovation has been expanding beyond the laboratory, some stakeholders may still lack the resources to engage with AI, with key themes of action, including:

1. Fostering intentional public engagement and public-private action to enhance sharing of best practices among all stakeholders
2. Increasing accessibility to responsibly curated AI-ready data, models and algorithms, and tooling and infrastructure for all

Below, HHS discusses the context of each theme of action in more detail, together with corresponding actions and plans to ensure equitable access to AI technologies and resources.

1. **Fostering intentional public engagement and public-private action to enhance sharing of best practices among all stakeholders:**

   Context:

   Increasing collaborative partnerships between stakeholders (e.g., the industry, STLTs, academia, and the general public) and intentional public engagement throughout the innovation pipeline could enhance the potential of AI being equitably adopted across medical research and discovery by sharing ideas, approaches, best practices, example applications, and key risks to mitigate between groups. HHS has already begun convening stakeholders and will continue to pursue actions to meet this challenge.

   HHS actions to date (non-exhaustive):

   - **NIH's AIM-AHEAD Program** seeks to build partnerships with underrepresented communities to develop and use AI in behavioral and biomedical research to establish networks to address health disparities.[171] This program spurs research and mentorship through projects that improve community engagement, leadership, and research fellowships (especially in underserved communities) and promote infrastructure development for AI in research.
   - NIH, NSF, NCSES, and DOE's **National AI Research Resource (NAIRR) Pilot** is a cross-agency collaboration working to improve AI in research, including research into topics related to human health.[172] It leverages large RWD sets to (1) build a synthetic data generator toolkit and framework to assess privacy risk and utility for using such data for evidence-building and (2) link medical imaging data with clinical records that will build capacity for multimodal AI development.

[171] https://datascience.nih.gov/artificial-intelligence/aim-ahead

[172] https://nairrpilot.org/

- **NIH's *All of Us* Research Program**[173] is a nationwide network of participant partners and researchers that aims to help ensure that people from all backgrounds can be included in research. Participants generously share information, which fuels thousands of studies to better understand health and disease, enabling more tailored and equitable approaches to care and creating new opportunities to leverage AI to advance precision medicine.
- **HHS is also developing challenges (i.e., innovation competitions), holding workshops (e.g., Evolving Landscape of Human Research with AI), and working with advisory committees to consult with members of the public to gather perspectives on tools** that facilitate data access, combination, and analysis (e.g., AI, cloud computing).[174, 175]

HHS near-term priorities:

- Promote and facilitate legal pathways for public-private partnerships (e.g., through the Foundation for the National Institutes of Health) between AI developers and NIH-funded investigators.
- Develop a vision and framework to incorporate public voices in all phases and types of clinical research.[176]
- Explore opportunities for public engagement and education in digestible forms about benefits, risks, and potential uses of AI in medical research and discovery to establish trust and promote uptake equitably.
- Continue to engage stakeholders (see Exhibit 3), including the public and participants, as part of the medical research and discovery pipeline to gather their perspectives on AI applications.
- Expand opportunities for collaboration and the implementation of initiatives for improving the AI readiness of NIH-supported data.[177]
- Facilitate public-private collaborations to foster AI knowledge and technology sharing by NIH-funded research institutions and underserved or underrepresented institutions.

HHS long-term priorities:

- Explore increasing resourcing for multi-institutional research collaborations, especially those embedding bioethicists and developers.
- Offer secure sandboxes[178] and infrastructure to encourage collaborative research into the development and use of AI for medical discovery, provided they ensure the development of information and communication technology (ICT) conforms to HHS Digital Accessibility Guidelines.[179]
- Facilitate community engagement, which will seed, sprout, and sustain long-term relationships between investigators and public members that can be utilized for co-creation. New authorities may be needed to survey stakeholders (including through AI, accounting for, or obtaining exemptions from constraints from the Paperwork Reduction Act). A new policy may be necessary to responsibly regulate such partnerships.

2. Increasing accessibility to responsibly curated AI-ready data, models and algorithms, and tooling and infrastructure for all:

(See Goal 1: "Catalyze Health AI Innovation and Adoption" theme of action 3: "Fostering AI-ready data standards and datasets to bolster their usability for AI-empowered medical research and discovery" for more information on data standards and usability)

---

[173] https://allofus.nih.gov/

[174] https://www.hhs.gov/ohrp/education-and-outreach/exploratory-workshop/2024-workshop/index.html.

[175] https://osp.od.nih.gov/policies/novel-and-exceptional-technology-and-research-advisory-committee-nextrac/, https://www.hhs.gov/ohrp/sachrp-committee/recommendations/irb-considerations-use-artificial-intelligence-human-subjects-research/index.html, https://www.hhs.gov/ohrp/sachrp-committee/recommendations/attachment-e-july-25-2022-letter/index.html NIH NExTRAC charges for data science and emerging technologies.

[176] https://osp.od.nih.gov/policies/novel-and-exceptional-technology-and-research-advisory-committee-nextrac This is the current charge of an NIH FACA called the NExTRAC.

[177] https://datascience.nih.gov/artificial-intelligence/initiatives/Improving-AI-readiness-of-Existing-Data

[178] See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan.

[179] https://www.hhs.gov/web/section-508/index.html

Context:

Effectively and efficiently harnessing AI requires financial, technical, and human resources. Though not a commodity, general-purpose AI technologies (e.g., LLMs) are widely available and will likely "raise the floor" of industrywide capabilities. The potential for more diverse researchers and use cases to apply these technologies in medical research and discovery could be hampered by resource availability, which could exacerbate an already prevalent "digital divide." HHS has made data and tools more accessible and plans to continue iterating on these activities.

HHS actions to date (non-exhaustive):

- **The NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative**[180] provides HHS-funded behavioral and biomedical investigators with discounted access to commercial cloud services, including AI applications. STRIDES has already generated approximately $120M in cost savings for these researchers, who can also access the associated "Cloud Lab," a sandbox[181] with associated tutorials and data where researchers can experiment with these technologies at no cost.
- **The NIH Policy for Data Management and Sharing** requires investigators to prospectively plan for maximizing appropriate sharing of "scientific data" (i.e., data of sufficient quality to validate and replicate research findings) and comply with the NIH-approved plan.[182] Supplemental information accompanying the policy helps researchers select a data repository, budget for data management and sharing, and protect human research participant data.[183, 184]
- **NIH's *All of Us* Research Program,**[185] also referenced above in the theme of action "fostering intentional public engagement and public-private action to enhance sharing of best practices among all stakeholders," is additionally building a diverse database that can inform thousands of studies on various health conditions. The program has created one of the largest, most diverse, and most broadly accessible health research datasets ever assembled. Data available to researchers include genomic data, survey responses, physical measurements, electronic health record information, and wearables data. The program's cloud-based platform design encourages collaboration across agencies, allowing researchers to leverage AI and related tools and expand their understanding of many health conditions.
- **ARPA-H's Biomedical Data Fabric Toolbox,**[186] in partnership with NIH, seeks to make it easier to connect biomedical research data from thousands of sources by (1) lowering barriers to high-fidelity, timely data collection in computer-readable forms, (2) preparing for multisource data analysis at scale, (3) advancing intuitive data exploration, (4) improving stakeholder access while maintaining privacy and security measures, and (5) ensuring generalizability of biomedical data fabric tools across disease types. These data must be findable, accessible, interoperable, and reusable.
- **NIH's Generalist Repository Ecosystem Initiative** supports seven generalist repositories that work together to establish consistent metadata, develop use cases for data sharing and reuse, and train and educate researchers on how to share and reuse data, including for the development and use of AI.[187]

HHS near-term priorities:

- Explore targeting research resources, training, and workshops to "expand the base" of AI-capable research institutions with a potential focus on data infrastructure.

---

180 https://datascience.nih.gov/strides

181 See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan.

182 https://sharing.nih.gov/

183 https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/data-sharing-approaches

184 https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/budgeting-for-data-management-sharing

185 https://allofus.nih.gov/protecting-data-and-privacy/precision-medicine-initiative-privacy-and-trust-principles

186 https://arpa-h.gov/research-and-funding/programs/arpa-h-bdf-toolbox

187 https://datascience.nih.gov/data-ecosystem/generalist-repository-ecosystem-initiative

- Explore resourcing for opportunities to continue supporting lower-resourced institutions to gain access to infrastructure (e.g., storage, computing, models) that is critical for AI adoption in medical research and discovery.
- Expand the availability and capability of resources like NAIRR, GREI, and ScHARe.
- Evaluate the expansion of the STRIDES program to include AI tools and models.
- Expand the availability, capability, and knowledge and tool/technology sharing from federal data initiatives.
- Develop as a public resource a federated, linked, centralized repository of AI-ready data for authorized stakeholders to engage in medical research and discovery.
- Continue developing data platforms that can be leveraged publicly to generate insights through AI that guide medical research and discovery.

HHS long-term priorities:

- Increase capacity to assist investigators in refining standards for data management and sharing in line with the changing landscape of public access to research.
- Build an internal database to track compliance, public comments, and other AI accessibility issues in medical research and discovery.[188]

1.6.4 Cultivate AI-Empowered Workforces and Organization Cultures

Without sufficient AI experts to enable innovation at scale in medical research and discovery, a widescale adoption and an uptake may be unfeasible. To that end, HHS plans to spur workforce development externally and internally to empower continued responsible, safe innovation of AI across the medical research and discovery value chain. Current themes of action in the space include:

1. Improving training in governance and management of AI in medical research and discovery
2. Developing and retaining a robust AI talent pipeline in medical research and discovery

Below, HHS's current actions and future goals to create AI-empowered workforces and organizational cultures in medical research and discovery are described.

1. **Improving training in the governance and management of AI in medical research and discovery:**

   Context:

   Most individuals involved in AI will be responsible for managing and using such technologies rather than developing them. Ensuring that the medical research and discovery enterprise gets the most out of AI will require focusing on the technologies and, perhaps more importantly, paying attention to their implementation, workflow integration, and life cycle management. Training the medical research and discovery workforce to manage and use such technologies responsibly will also be critical to harnessing AI to advance the industry. HHS has addressed this challenge and will direct additional efforts to resolve this gap further and empower the industry.

   HHS actions to date (non-exhaustive):

   - **FDA's blog entry, "A Lifecycle Management Approach Toward Delivering Safe, Effective AI-Enabled Healthcare,"**[189] provided an overview of one potential approach to developing, validating, and maintaining ongoing governance of AI models for medical devices to ensure their safety and effectiveness.

[188] https://www.consumerfinance.gov/data-research/consumer-complaints/

[189] https://www.fda.gov/medical-devices/digital-health-center-excellence/blog-lifecycle-management-approach-toward-delivering-safe-effective-ai-enabled-health-care

HHS near-term priorities:

- Explore targeting resources, training, and workshops to include the governance, management, and use of AI technologies in research and technology.
- Consider supporting guidelines or best practices for governance, life cycle management, and workflow integration of AI technologies in medical research and discovery.

HHS long-term priorities:

- Iteratively amend and publish updates to guidelines or training programs as appropriate.

2. **Developing and retaining a robust AI talent pipeline in medical research and discovery:**

Context:

To harness the potential of AI in medical research and discovery, the ecosystem may need a strong and diverse workforce pipeline capable of integrating models and algorithms into their inquiries. Different types of AI are likely to shift the skillsets and roles needed for an effective medical research and discovery workforce as multimodal models become increasingly powerful and potentially automate many aspects of the scientific workflow (from observation and hypothesis development to data analysis and manuscript development), human input and evaluation will be necessary at all stages. Investigators from all backgrounds may need baseline knowledge to develop and apply AI safely, responsibly, and effectively. Additionally, without clear incentives, interdisciplinary experts may continue to flow toward the technology sector, leaving gaps in non-profit, academic, and government laboratories focused on medical research and discovery. HHS has taken action to meet this challenge and plans to continue exploring opportunities.

HHS actions to date (non-exhaustive):

- **NIH's AIM-AHEAD Program** established a strong mentoring network to cultivate AI talent in medical research and discovery across the U.S.[190]
- **The NIH DATA National Service Scholar Program** hired data science professionals to NIH to increase efficiency, innovative research, tool development, and analytics in research.[191]
- **NIH's Administrative Supplements for Workforce Development at the Interface of Information Sciences, AI, and Biomedical Sciences** supports the development and implementation of curricular or training activities at the interface of information science, AI, and biomedical sciences to develop the competencies and skills needed to make biomedical data findable, accessible, interoperable, and reusable and AI-ready.[192]
- **National Library of Medicine's (NLM's)**[193] **University-based Biomedical Informatics and Data Science Research Training Programs** support research training in biomedical informatics and data science at graduate and post-doctoral educational institutions in the U.S.[194]
- **NLM's Short-Term Research Education Experiences to Attract Talented Students to Biomedical Informatics/Data Science Careers and Enhance Diversity** supports educational activities that encourage talented undergraduate and master's students, including those from groups underrepresented in the biomedical and behavioral sciences, to pursue further training and careers in biomedical informatics and data science. NLM seeks to develop a cadre of diverse scientists capable of leading biomedical informatics and data science research with this program.[195]

[190] https://datascience.nih.gov/artificial-intelligence/aim-ahead

[191] https://datascience.nih.gov/data-scholars-2023

[192] https://datascience.nih.gov/artificial-intelligence/initiatives/Workforce-Gap-Data-Governance-AI

[193] Note that NLM is a subsidiary of NIH.

[194] https://www.nlm.nih.gov/ep/GrantTrainInstitute.html

[195] https://www.nlm.nih.gov/ep/R25_program.html

- **NLM's Data Science and Informatics (DSI) Scholars Program** is an 8- to 12-week summer internship in which interns contribute their skills and perspectives to computational research projects in the biological sciences. DSI Scholars gain valuable experience in a collaborative research environment while training one-on-one with a research mentor.[196]

HHS near-term priorities:

- Prioritize and explore resourcing for evidence-building to evaluate AI workforce development efforts and maximize the efficacy of HHS spending.
- Increase and amplify training for researchers on developing responsible AI tools for medical research and discovery, including best practices for integrating AI-related coursework into biomedical research training curricula.
- Integrate biosecurity resources or training to share with researchers new to utilizing AI.
- Create education and training programs for providers on the use of AI in medical research and discovery and how patient data can be used and collected to propel further innovation safely.
- Evaluate the expansion of **NIH's AIM-AHEAD Program** to include recruitment and training for AI expertise in medical research and discovery.

HHS long-term priorities:

- Explore expanding resourcing mechanisms that emphasize the development and use of AI in biomedical research graduate training.
- Explore resourcing for centers of excellence for data science and AI in research institutions across the U.S. that offer subsidized training and services for HHS-funded researchers.
- Promote community-driven training for upskilling in prompt engineering, red teaming, and watermarking to maximize the utility of AI while maintaining scientific rigor and driving equity.

1.7 Conclusion

Fostering innovation while managing risks in AI-driven medical research and discovery is crucial for advancing American health and human services. HHS understands that the potential of AI to enhance research outcomes, speed up the development of medical products, and improve patient care is vast; however, these benefits must be balanced against the risks of bias, data misuse, biosecurity, and other concerns. HHS is uniquely positioned to play a pivotal role in this landscape. HHS's action plan—which includes initiatives exploring resourcing, public education, and workforce development—aims to address current challenges to AI adoption in medical research and discovery and advancing its safe and responsible use. By doing so, HHS can stimulate economic growth, create high-skilled jobs, and, most importantly, safeguard the health and well-being of all Americans and individuals globally. Through strategic leadership and collaboration with stakeholders across the value chain, HHS can guide the responsible integration of AI in medical research and discovery, helping to ensure that the benefits of innovation are realized while associated risks are mitigated. HHS is committed to evolving its AI strategy in medical research and discovery as technologies and use cases continuously change to best improve medical research and discovery.

[196] https://www.nlm.nih.gov/research/DDSI.html

2 Medical Product Development, Safety, and Effectiveness

2.1 Introduction and Context

Medical products, including drugs,[197] biological products,[198] and medical devices,[199] including some software-based behavioral interventions,[200] play a crucial role in advancing health. As AI becomes increasingly advanced, it has the potential to further improve patient care by augmenting the capabilities of healthcare practitioners and bolstering product development across the life cycle from clinical trials to manufacturing and safety monitoring.[201] The rapid advancement of AI technologies in the medical products space places HHS in a pivotal position. HHS can spur the successful adoption and scale-up of effective technologies while minimizing potential risks and harm associated with medical products throughout their life cycle.[202]

This chapter of the Plan will focus on medical products themselves and steps of the medical product lifecycle from clinical trials to regulatory review, manufacturing, and safety monitoring. For more information on the research and discovery of medical products[203] and the research and discovery of AI technologies that can be leveraged in biomedicine, please refer to the Medical Research and Discovery chapter.

The role of AI in devices differs from other medical products. In drugs and biological products, it is generally helpful in producing information or data to support decision-making across the product development life cycle, from development to manufacturing and postmarket surveillance and monitoring. In devices, it may play three roles: in the development or maintenance of the device, as a stand-alone product that can perform one or more device purposes (e.g., diagnose, cure, mitigate, treat, or prevent disease) without being a part of a traditional hardware device, or as part of or integral to a device.

Regulatory review for marketing authorization of these products in the U.S. is governed by a statutory and regulatory framework that helps ensure medical products are safe and effective for their intended use. Across the product life cycle, FDA reviews data and information about products before they are marketed to the public, conducts surveillance once products are available, and monitors product promotion and medical product quality.[204]

As of August 2024, FDA has authorized approximately 1,000 AI-enabled medical devices,[205] and FDA has received over 550 submissions for drug and biological products with AI components.[206] NIH also plays a critical role in advancing the development of medical products that increase access to better care. Though funding for clinical development can come from a variety of places, NIH alone makes an approximately $3B annual

[197] See Appendix A: "Glossary of terms" for the definition of "drug" used in this Plan.

[198] See Appendix A: "Glossary of terms" for the definition of "biological product" used in this Plan.

[199] See Appendix A: "Glossary of terms" for the definition of "medical device" used in this Plan.

[200] Some software-based behavioral interventions are medical devices under FDA's statute, whereas others, such as those software functions that are "intended for maintaining or encouraging a healthy lifestyle" and are "unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition" are not. See sections 201(h) and 520(o)(1)(B) of the FD&C Act.

[201] https://www.fda.gov/media/177030/download

[202] https://www.hhs.gov/programs/topic-sites/ai/strategy/index.html

[203] Drugs, biological products, and medical devices in this Plan are referred to as "medical products" when discussed collectively. See Appendix A: "Glossary of terms" for the definition of "medical products" used in this Plan for additional details.

[204] https://www.fda.gov/patients/learn-about-drug-and-device-approvals

[205] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

[206] https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/artificial-intelligence-drug-development

investment in clinical trials, making it the largest federal funder of clinical trials in the U.S.[207] Regulatory oversight of medical products strives to maintain a balance between upholding safety and effectiveness and fostering innovation, including when AI is used in the medical product or across the medical product life cycle.

2.1.1 Action Plan Summary

Later in this chapter, HHS articulates proposed actions to advance its four goals for the responsible use of AI in the sector. Below is a summary of the themes of actions within each goal. For full details of proposed actions please see section 2.6 Action Plan.

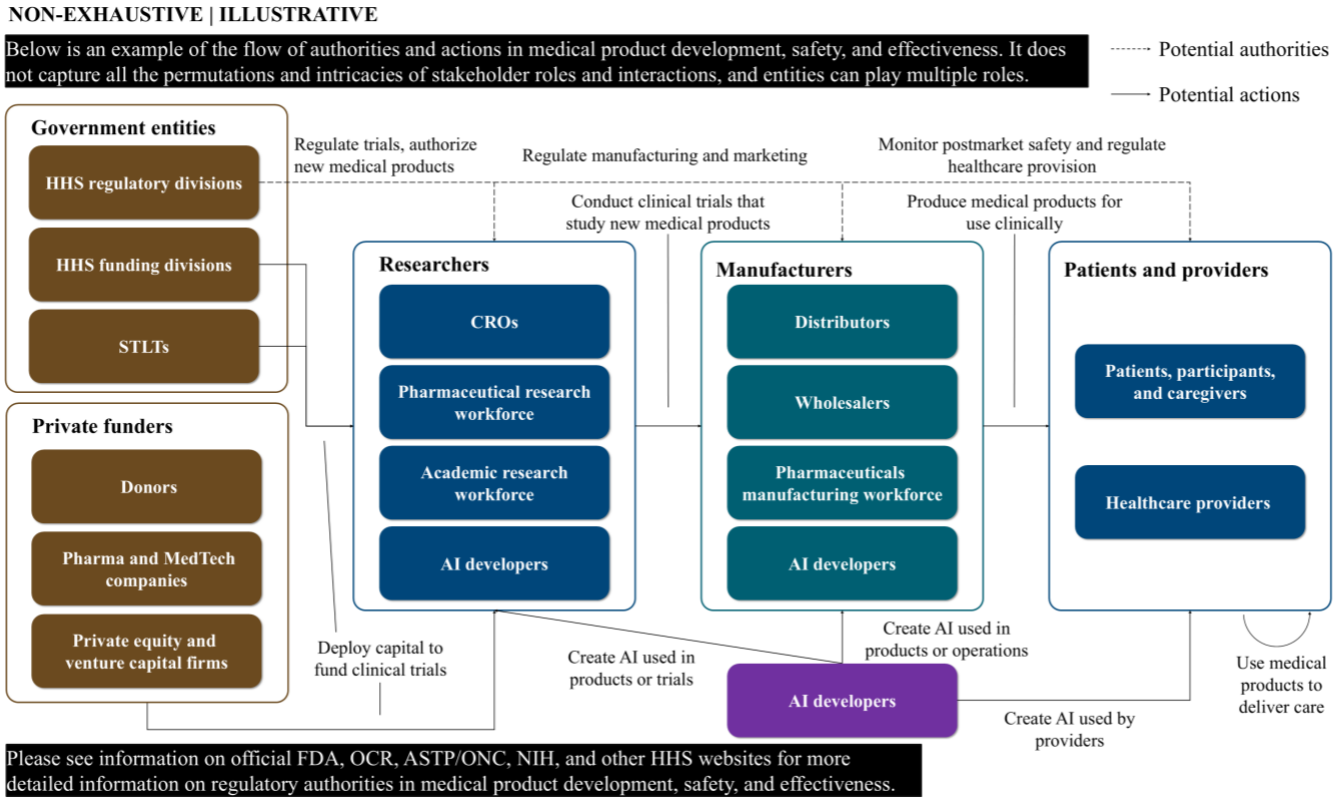
Key goals that actions support	**Themes of proposed actions *(not exhaustive, see 2.6 Action Plan for more details)***
1. Catalyzing health AI innovation and adoption	• Clarifying regulatory oversight of medical products • Providing clarity on payment models • Fostering public-private partnerships and intergovernmental collaborations to rapidly develop and share knowledge
2. Promoting trustworthy AI development and ethical and responsible use	• Refining regulatory frameworks to address adaptive AI technologies in medical devices • Promoting equity in AI deployment to bolster safe and responsible use • Addressing AI-enabled software outside current device regulatory authorities • Fostering private or public mechanisms for quality assurance of health AI
3. Democratizing AI technologies and resources	• Enabling collaborative development through public engagement • Aligning standards and information-sharing mechanisms across research and healthcare delivery
4. Cultivating AI-empowered workforces and organization cultures	• Improving training in the governance and management of AI in medical products • Developing and retaining AI talent related to medical products

2.2 Stakeholders Engaged in Medical Product Development, Safety, and Effectiveness

A range of stakeholders engage with AI in medical products and their development, ranging from patients and medical providers to developers of medical products, distributors, providers, payers, researchers, and many others. The Action Plan section at the end of this chapter includes approaches to engage these stakeholders to advance innovation while mitigating risks. Below is an illustrative diagram of example flows between stakeholders and a bulleted list with additional details on stakeholders involved in medical product development, safety, and effectiveness. Please note that neither the diagram nor the list captures all possible stakeholder roles and interactions. Please refer to other HHS documents for additional details on regulatory guidance and authorities.

[207] https://grants.nih.gov/policy-and-compliance/policy-topics/clinical-trials/why-changes

Exhibit 5: Stakeholders Engaged in Medical Product Development, Safety, and Effectiveness



Stakeholders include, but are not limited to:

- **HHS operating divisions (non-exhaustive):**[208]
  - **FDA:** Helps ensure that human and animal drugs, biological products, and medical devices are safe and effective for their intended uses and that electronic products that emit radiation are safe. As AI becomes a more prominent aspect of medical products, their development, manufacturing operations, and use, FDA will continue to regulate and support stakeholders.
  - **NIH:** Supports biomedical and behavioral research within the U.S. and abroad, conducts research in its laboratories and clinics, trains promising young researchers, and promotes collecting and sharing biomedical knowledge, which have increasingly included AI related to medical products and the life cycle.
  - **CDC:** Develops recommendations on using vaccines after the FDA approves them, continually monitors vaccines for safety once used clinically, and reports adverse effects (e.g., via the Vaccine Adverse Event Reporting System).[209]
  - **AHRQ:** Supports research on interventions enabled by medical devices, such as patient-centered clinical decision support, and focuses on improving the quality, safety, efficiency, and effectiveness of healthcare for all Americans.
- **Other federal agencies:** HHS also works closely with many other federal departments, such as the National Science Foundation (NSF) and the Department of Energy (DOE).
- **Patients, participants, and caregivers (including residents and communities):** Use drugs, biological products, or medical devices developed using AI or including AI. Today, empowered patients may also utilize AI to better understand their personal health status and advocate for their own care.

[208] https://www.hhs.gov/about/agencies/hhs-agencies-and-offices/index.html

[209] https://www.cdc.gov/vaccines-children/about/developing-safe-effective-vaccines.html

- **Pharmaceutical and medical technology research and manufacturing companies:** Design, develop, and produce drugs, biological products, or medical devices for commercial use in healthcare delivery, including researchers and subject matter experts integrating AI into clinical trials and product design and manufacturing. They are among the primary users of AI in clinical trials and medical product manufacturing. These companies also use AI to support pharmacovigilance activities.
- **Healthcare providers and payers:** Utilize medical products and provide clinical perspectives to clinical development efforts (e.g., hospitals, clinics, healthcare professionals) or decide which technologies are part of its payment mechanisms (e.g., payers). Additionally, providers can be "humans in the loop" for AI use, which includes portions of the medical product life cycle. The use of AI in clinical settings is expanded on in the Healthcare Delivery chapter, as medical product use intended by manufacturers and authorized by the FDA could be leveraged to provide healthcare for certain purposes while not changing their device, drug, or biological product status.
- **STLTs:** Play oversight and funding roles outside of the federal government. FDA has regulatory oversight of medical products, while STLTs may have jurisdiction over different components of medical practice and healthcare delivery.
- **Academic, non-profit, and other research workforce:** Develop evidence for the leading edge of biomedical knowledge, including engineers designing and generating medical devices for clinical applications, and subject matter experts developing AI, applying AI in clinical trial workflows, and/or integrating AI into the product development life cycle. They are among the primary users of AI in medical product development.
- **Contract research organizations (CROs):** Provide outsourced research services and may develop or integrate AI into their clinical trial workflows. As third parties, CROs should be engaged particularly on matters of security and privacy as they handle other organizations' sensitive data in AI. AI is also used by these companies to support pharmacovigilance activities that may be outsourced by drug manufacturers.
- **Distributors and wholesalers:** Facilitate the distribution of medical products—which may include or have been researched and developed leveraging AI—to healthcare providers.
- **Donors and private funders:** Support funding for product development and scale-up. They include non-profit donors, such as foundations, and for-profit funders, such as private equity, venture capital, and other funding organizations. These organizations may also support other investments in AI technologies or with other stakeholders.
- **AI developers:** Build the AI tools, models, and platforms that can be used within medical products or across the medical product life cycle.

2.3 Opportunities for the Application of AI in Medical Product Development, Safety, and Effectiveness

If adopted and scaled successfully and responsibly, AI use in medical product development, operations, and safety monitoring, as well as AI inclusion in the medical product itself, could improve overall care outcomes and the accessibility and efficiency of the process in multiple ways, such as:

1. **Increasing the efficiency of clinical trials, which may accelerate the timeline to access safe and effective medical products:** Leveraging AI in clinical trials may help predict a participant's risk for adverse reactions, generate initial content of regulatory submissions and investigative brochures, and translate documentation to other languages. Additionally, though there are methods to incorporate patient centricity without AI, using AI toward this goal may reduce the likelihood of candidate attrition.[210] Furthermore, using AI to execute analyses can accelerate another core part of the clinical trial process. Together, these and other

[210] https://pmc.ncbi.nlm.nih.gov/articles/PMC11006977/

uses of AI in clinical trials can make medical products accessible to patients more rapidly. (See trend (A)(1) in the section below for more details on AI uptake in clinical trials to date).

2. **Improving the representativeness of clinical trials of those who use medical products:** Today, as many as "86% of clinical trials do not reach recruitment targets within their specified time periods,"[211] which can lead to less effective medical interventions, potentially poorer health equity in pharmaceutical practices, and potentially billions of dollars in economic losses.[212] Leveraging AI in clinical trial strategy, as appropriate, to analyze patient and other demographic data, to select sites, and to identify potential candidates that are representative of the population of interest has the potential to help enroll a more representative population in clinical trials. This can bolster the information submitted to the FDA for regulatory approval or marketing authorization. Leveraging AI in clinical trial strategy can better serve historically underrepresented populations.
3. **Being used as part of a medical product, being the medical product itself, or being used to develop medical products:** AI can be used as part of a medical product or to develop safe and effective medical products. In particular, AI-enabled medical devices, such as over-the-counter hearing aids, have the potential to be used by patients, healthcare providers, and other end users to help augment care and improve outcomes.[213, 214] (See trend (B)(1) in the section below for more details on AI-enabled medical devices). Additionally, AI supports the ability to learn from data collected during clinical use which can help support improving medical product accuracy and performance over time,[215] potentially leading to improved accuracy and monitoring (e.g., lower misdiagnosis rates, higher ability to detect adverse effects early). Similarly, AI can be leveraged to develop drugs and biological products (e.g., identifying targets and assessing biomarkers and endpoints) as discussed in the Medical Research and Discovery chapter.
4. **Strengthening supply chain, manufacturing, and other operations to ensure and expand access:** In recent years, medical product supply shortages have impacted patients' ability to access timely care that is critical for their health. For example, as of October 2024, there are over 100 active drug shortages, spanning from IV solutions to prescription stimulants.[216] Similarly, when demand for a specific medical product surges, increasing access by rapidly driving up supply may not be a quick process.[217] AI can rationalize and streamline supply chain management and manufacturing processes, including the ability to analyze production schedules, forecast demand, estimate the impact of potential disruptions, and optimize inventory.[218] By responsibly adopting AI into their operations, medical product manufacturers, distributors, and others can mitigate shortages, safeguard access to care, and prepare for expansion to additional patients when demand spikes.
5. **Enhancing pharmacovigilance and postmarket surveillance and monitoring:** Monitoring medical products is crucial to managing their safe and effective use. Today, data collection and analysis already leverage EHRs, administrative claims, and other sources of clinical data to collate large amounts of product safety data (e.g., FDA's Adverse Event Reporting System [FAERS] and FDA's Sentinel Initiative).[219, 220] Some safety monitoring activities involve surveys and social media monitoring, which can take substantial resources and time.[221] Leveraging AI to collect and/or analyze large datasets of adverse event reports, scraped social media data, or survey data could rapidly identify potential safety issues and accelerate the timeline for taking action to protect patients. Furthermore, this data and analysis could be leveraged to better understand the outcomes of medical product use and derive novel insights to enhance human health,

[211] https://www.sciencedirect.com/science/article/pii/S155171441730753X#bb0020
[212] https://www.ncbi.nlm.nih.gov/books/NBK584396/
[213] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device
[214] https://www.fda.gov/news-events/press-announcements/fda-authorizes-first-over-counter-hearing-aid-software
[215] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device
[216] https://www.drugs.com/drug-shortages/
[217] https:/www.ncbi.nlm.nih.gov/books/NBK583734
[218] https://www.fda.gov/media/167973/download?attachment
[219] https://www.fda.gov/drugs/fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard
[220] https://www.fda.gov/safety/fdas-sentinel-initiative
[221] https://www.nsf.org/knowledge-library/post-market-surveillance-what-you-need-to-know-to-ensure-patient-safety

including the types of patients best served by a particular medical product. One caution, however, is that with potentially large quantities of clinical data, more noise could be generated, so parsing essential signals from the data is paramount.[222]

2.4 Trends in AI in Medical Product Development, Safety, and Effectiveness

Stakeholders have begun to leverage AI in medical products and their development along two overarching trends:

A. Leveraging AI in the development of medical products and their ongoing operations
B. Embedding AI within products themselves or as standalone products

Below, select non-exhaustive examples of adoption across (A) and (B) to date are discussed.

A. **AI in the development and operations of medical products**
    1. **AI uptake related to drugs and biological product development is increasing:** There has been a growing use of AI in the drug and biological product development life cycle across a range of TAs. In fact, FDA has seen a significant increase in the number of drug and biological product application submissions using AI components over the past few years, from just 3 in 2018 to 132 in 2021.[223, 224] These submissions traverse the landscape of drug and biological product development ranging from clinical research to postmarket surveillance and monitoring and advanced pharmaceutical manufacturing.[225] Use cases seen in recent FDA submissions focused on a range of topics, including but not limited to endpoint and biomarker assessment, anomaly detection, imaging, video, and voice analysis, patient risk stratification and management, dose selection and optimization, and adherence during clinical trials.[226] Additional use cases span some of the most time-intensive aspects of clinical trials (e.g., site selection and candidate recruitment) and can help predict the success or failure of proposed trial designs.[227] AI is also being leveraged to reduce the time associated with and to increase the quality of randomized controlled trials by selecting participants and minimizing errors.[228]
    2. **Approaches to validate the credibility of health AI are heterogenous and inconsistently applied:** The use of AI in the health domain, including in the development and operations of medical products, needs to be validated to ensure that it leads to safe and effective medical products, decisions, and actions. Today, there are many AI validation approaches, and in general, they focus on easy-to-measure quantitative performance metrics in narrow and highly controlled conditions and rarely use real patient data.[229] The ease with which AI can be deployed to a wide and ever-expanding array of healthcare use cases is driving a potential need to establish nationally accepted standards and mechanisms for assuring the quality of AI systems.

B. **AI within or as the products**
    1. **Applications of AI-enabled medical devices are expanding, with a focus on radiology:** Within medical devices, AI has grown rapidly to cover new applications across the medical product ecosystem. As of August 2024, the FDA has reviewed and authorized approximately 1,000 AI-enabled medical devices to market in the U.S.,[230] including 171 in 2023 and 258 in 2024,[231, 232]

[222] https://psnet.ahrq.gov/perspective/artificial-intelligence-and-patient-safety-promise-and-challenges
[223] https://ascpt.onlinelibrary.wiley.com/doi/10.1002/cpt.2668
[224] https://www.fda.gov/news-events/fda-voices/harnessing-potential-artificial-intelligence
[225] https://www.fda.gov/media/167973/download?attachment
[226] https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/cpt.2668
[227] https://www.nature.com/articles/d41586-024-00753-x
[228] https://pmc.ncbi.nlm.nih.gov/articles/PMC7346875/
[229] https://pubmed.ncbi.nlm.nih.gov/39405325/
[230] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices
[231] https://rad.washington.edu/news/fda-publishes-list-of-ai-enabled-medical-devices/
[232] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

which indicate a 33% and 27% increase in authorized AI-enabled medical devices in the last two years, respectively.[233] Over 75% of these devices are used in a radiology context, potentially due to the high number of predicate devices that may enable clearer paths to 510(k) clearance. Additionally, FDA-authorized devices use predictive AI rather than GenAI, which is more nascent.

See trend (A)(1) for trends in drugs and biological products clinical development, the "Table of Example Use Cases and Risks Across Steps of the Medical Product Life Cycle That Are in the Scope of This Chapter," which follows for use cases in drugs and biological products clinical development, and the Medical Research and Discovery chapter generally for trends and use cases of AI in drugs and biological products discovery, which are potentially the most prevalent and mature areas of uptake.

2.5 Potential Use Cases and Risks for AI in Medical Products and Their Development

Below, parts of the medical product life cycle that are in the scope of this chapter are described similarly to the "value chains" outlined in other chapters in this Plan to help guide the subsequent discussion on use cases and risks. Note that pre-clinical steps of the medical product life cycle (e.g., basic research, discovery) are discussed in the Medical Research and Discovery chapter.

Exhibit 6: Steps of the Medical Product Life Cycle That Are in the Scope of This Chapter

The above diagram showcases the overarching medical product life cycle in the scope of this chapter, from clinical development to monitoring product safety postmarket.

Development processes for drug and biological products and medical devices follow the same overarching steps, though processes differ within those steps, particularly in regulatory approval. Each step of the medical product life cycle shown in the exhibit above is explained below:

1. **Clinical development** differs between drugs and devices as summarized below:
    a. *Drugs and biological products*: Before a clinical trial with a drug or biological product can proceed, an Investigational New Drug (IND) application for drugs and biological products must be submitted to the FDA.[234] At a high level, drug development involves a series of clinical studies with human subjects to assess the safety and effectiveness of candidate technologies, generally divided into three phases: Phase I tests safety and dosage. Phase II evaluates preliminary efficacy and safety, and Phase III further evaluates efficacy and safety.[235] In certain cases, such as with certain vaccines or drugs, the FDA may require a Phase IV trial or postmarket safety study to assess known or potential serious risks further.[236]
    b. *Medical devices*: The device development program does not typically follow the same drug phasing sequence. If a particular device does require testing in clinical trials prior to FDA marketing authorization, it may require an investigational device exemption (IDE),[237] although many software-

[233] Only through August 2024, potentially higher by the end of the full year
[234] https://www.fda.gov/drugs/types-applications/investigational-new-drug-ind-application
[235] https://www.fda.gov/patients/drug-development-process/step-3-clinical-research
[236] https://www.fda.gov/vaccines-blood-biologics/development-approval-process-cber/vaccine-development-101
[237] https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/investigational-device-exemption-ide

based device studies are not significant risk and proceed under the oversight of an institutional review board (IRB) only.[238]

2. **Regulatory review** may differ for drugs and biological products versus medical devices. Given the complexities of review processes, this Plan will not attempt to summarize steps but rather point to FDA's resources on both below:
    a. *Drugs and biological products*: A detailed description of the development and approval process for drugs and biological products can accessed in the footnotes.[239, 240]
    b. *Medical devices*: A detailed description of the marketing authorization process for medical devices can be accessed in the footnotes.[241]
3. **Manufacturing and supply chain** refers to the operational process of procuring necessary materials, using them to develop medical products, and distributing them downstream to customers after a product has marketing authorization. Manufacturers must comply with applicable regulatory requirements, which include FDA's Quality System Regulation/Medical Device Current Good Manufacturing Practices (CGMP)[242] and drug and biological product CGMP regulations,[243, 244, 245] which assures that medical products are not adulterated during production.
4. **Market access, commercial, and other operations** involve developing and distributing materials that explain the relevance and impact of the product if leveraged in various care situations for potential providers, payers, or other stakeholders. These include logistics, sales, pricing, finance, health, economics, outcomes research, and other enabling stakeholder activities. FDA regulates the marketing of medical products, including but not limited to preventing false or misleading labeling of medical products.[246]
5. **Postmarket monitoring for safety and effectiveness** includes using medical products in clinical settings, consistently monitoring their safety, and identifying and mitigating issues to ensure ongoing patient safety. Requirements for the postmarket monitoring of medical devices include reporting device malfunctions, serious injuries or deaths, and inspecting establishments where devices are produced or distributed.[247] With respect to drugs, the FDA carefully monitors performance through FAERS and the Sentinel Initiative.[248, 249] Additionally, vaccines, in particular, are closely monitored via various surveillance systems, such as the Vaccine Adverse Event Reporting System, the FDA BEST (Biologics Effectiveness and Safety) program, and the CDC's Vaccine Safety Datalink.[250]

AI uptake has tremendous potential to drive innovation in medical products and across the medical product life cycle to benefit patients, which should be implemented with careful attention to risk mitigation.

While risks differ between AI related to drugs, biological products, and devices, a few high-level themes emerge that could be important to consider as technology rapidly advances. In clinical development, AI can perpetuate biases inherent in the data on which it was trained or tuned. As part of manufacturing and supply chain, when using AI for tracking and managing the supply chain for manufacturing, potential risk may arise from inaccuracies in AI projections of supply needs, leading to insufficient production. Insufficient production may lead to shortages, leaving people without access to medical products critical to their care. Given these themes and other

[238] https://www.fda.gov/medical-devices/investigational-device-exemption-ide/ide-institutional-review-boards-irb
[239] https://www.fda.gov/drugs/development-approval-process-drugs
[240] https://www.fda.gov/vaccines-blood-biologics/development-approval-process-cber
[241] https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/how-study-and-market-your-device
[242] https://www.fda.gov/medical-devices/postmarket-requirements-devices/quality-system-qs-regulationmedical-device-current-good-manufacturing-practices-cgmp
[243] https://www.ecfr.gov/current/title-21/chapter-I/subchapter-C/part-210
[244] https://www.ecfr.gov/current/title-21/chapter-I/subchapter-C/part-211
[245] https://www.ecfr.gov/current/title-21/chapter-I/subchapter-F/part-600
[246] https://www.fda.gov/medical-devices/overview-device-regulation/device-labeling
[247] https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/postmarket-requirements-devices
[248] https://www.fda.gov/drugs/fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard
[249] https://www.fda.gov/safety/fdas-sentinel-initiative
[250] https://www.fda.gov/vaccines-blood-biologics/development-approval-process-cber/vaccine-development-101

risks described below, HHS is already working to safeguard against these risks and will continue to explore potential actions to encourage safe, innovative AI adoption in the space.

2.5.1 Table of Example Use Cases and Risks Across Steps of the Medical Product Life Cycle That Are in the Scope of This Chapter

AI is being adopted across the medical product life cycle. In the tables below, HHS highlights a non-exhaustive list of potential benefits, uses, and risks across the steps that are in the scope of this chapter as described above. Parties should consider applicable statutory and regulatory requirements and consult relevant regulatory agencies when appropriate. Please note that the use cases detailed below highlight existing or potential ways that AI can be used by a variety of stakeholders in this domain. For details on how HHS and its divisions are using AI, please reference the HHS AI Use Case Inventory 2024.[251]

Functional component 1: Clinical development

Includes studies with human participants to assess the safety and effectiveness of investigational medical products

Please note that the Medical Research and Discovery chapter discusses basic research and pre-clinical development, which includes a discussion on use cases and risks of AI related to target identification, lead and hit generation and optimization.

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive and analytical models that can help improve the representativeness of the trial population *E.g., site selection to maximize meeting enrollment goals* Helping to identify clinical study sites with representative patients to help meet enrollment goals[252] *E.g., candidate selection to help ensure a representative trial population* Leveraging advanced analytics to identify cohorts that are representative of the population that will use a product if approved[253]	**Potential to misdirect the course of research** *E.g., "false positives" or "false negatives" in clinical trials* In technology that augments researchers in clinical trials, AI could identify safety events that are not true events or fail to identify serious safety events. If the researcher relies too heavily on the AI characterization or makes a human error in oversight of the AI, this may lead to misclassification and impact the ability to draw conclusions when analyzing data. **Potential for bias** *E.g., lack of representativeness of population using a medical product* While AI can advance medical product development by identifying participants, designing trials, analyzing outputs, and more, it may not be trained on data representing the population that
Generative, representational, and predictive models that accelerate the timeline of clinical trials *E.g., strategy for clinical trials design that increases the probability of success by reducing the likelihood of rework* Leveraging generative and analytical models that can simulate potential trial designs and recommend a subset with the highest probability of success[254] *E.g., digital twins for faster, in silico experimentation* Representing objects, systems, or candidates virtually can accelerate research by enabling simulated testing of products[255]	

[251] https://www.healthit.gov/hhs-ai-usecases

[252] https://www.fda.gov/drugs/news-events-human-drugs/role-artificial-intelligence-clinical-trial-design-and-research-dr-elzarrad

[253] https://www.fda.gov/drugs/news-events-human-drugs/role-artificial-intelligence-clinical-trial-design-and-research-dr-elzarrad

[254] https://www.fda.gov/media/167973/download

[255] https://datascience.nih.gov/tools-and-analytics/digital-twins

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
E.g., endpoint assessment and biomarker identification Using AI as part of a clinical outcome assessment or to identify biomarkers that can potentially serve as endpoints in clinical trials[256] *E.g., image, video, and voice analysis to accelerate analyses and potentially bolster their quality* Leveraging AI, "usually deep learning, for the analyses of imaging data, videos, or voices" can contribute to faster and potentially more precise analyses[257] *E.g., patient risk stratification and dosage optimization to improve trial participant safety* Predicting dosages and patients' risk for a specific severe adverse event "based on patient baseline information" and subsequently using this prediction "to help determine the need of inpatient or outpatient monitoring for each patient"[258]	may ultimately use the medical product clinically. This could lead to research outcomes that are only relevant for a small group and potentially miss opportunities to address health disparities if AI models are not trained on representative data.

Functional component 2: Regulatory review

Submission of documents to the FDA

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Leveraging generative models to accelerate the development and enhance the quality of medical writing *E.g., auto-writing of clinical study reports (CSRs) to reduce researcher time spent drafting results* Leveraging natural language processing (NLP) and ML algorithms to synthesize results that could be included in regulatory submissions to the FDA when appropriately confirmed by humans[259] *E.g., the generation of medical content across all documents that could be submitted for regulatory approval* Generating first drafts of research or other medical documents from existing materials to increase the speed of document development and potentially bolster their quality when appropriately confirmed by humans[260]	**Potential for inaccuracies that lower chances of approval** *E.g., misaligned syntheses of patient or candidate records and healthcare professional (HCP) or researcher notes* Leveraging AI to synthesize or generate content related to patient records, sometimes with human-written notes involved, can lead to outputs that do not apply to the situation at hand because poor data quality can lead to poor outputs. Using such tools could require careful oversight regarding the types of data it analyzes and its output. **Potential to introduce safety risks** *E.g., generating insights from research results in regulatory submissions that are not based on data* Content generated by some AI (e.g., LLMs) can be inferred rather than based on facts, leading to regulatory submissions that contain inaccurate information. If not caught, such inaccuracies can lead to marketing authorizations for medical products that are not safe and effective.

[256] https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/cpt.2668
[257] https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/cpt.2668
[258] https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/cpt.2668
[259] https://pmc.ncbi.nlm.nih.gov/articles/PMC10492634/
[260] https://pmc.ncbi.nlm.nih.gov/articles/PMC10492634/

Functional component 3: Manufacturing and supply chain

Operations related to procurement, development of products, and distribution of those products downstream to customers

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive and monitoring tools that enable advanced identification of problems or inefficiencies *E.g., monitoring of manufacturing operations for real-time analysis and recommendations of actions to enhance operations* Receiving real-time data on drug production processes to improve productivity, correct inefficiencies, control quality, and predict yields[261]	**Potential to disrupt the supply of critical medical products** *E.g., disruptions to operations of critical drugs, biological products, and devices from AI-empowered monitoring of supply chain and manufacturing operations* If not properly implemented and managed with expert human oversight, using AI to track and manage the supply chain for raw materials can result in inaccuracies in AI projections of supply needs, leading to insufficient production. Insufficient production may lead to shortages, leaving people without access to medical products critical to their care.
Optimization algorithms that help to ensure patient needs are met, and the likelihood of shortages or product waste is reduced *E.g., maximization of production output of existing physical and operational infrastructure* Predicting the performance of operations, people, and machinery with automated inventory tracking to mitigate stockouts and supply delay risks[262]	**Potential for bias** *E.g., inequitable allocation of medical product supply* Leveraging AI to plan demand, logistics, and production for drug and medical device needs could result in disparate allocations if data used in AI analysis is not sufficiently representative of the population of patients served by the corresponding products. This could perpetuate existing health inequities and reinforce biases if impacted populations receive less access to the drugs, biological products, and devices needed for their health.

Functional component 4: Market access, commercial, and other operations

Connecting to potential healthcare providers and payers to explain the relevance and impact of medical products (includes pricing, finance, logistics, and enabling activities)

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Analytical and generative tools that streamline and bolster market entry activities *E.g., co-pilots for patient and HCP representatives to reduce knowledge gaps* Leveraging GenAI trained on details about all products to help answer questions quickly about topics patient and HCP representatives may be unfamiliar with *E.g., identification of inaccurate information in marketing materials* Using advanced analytics to scour the internet and other resources that promote medical products to compare against FDA-approved labeling and flag potential regulatory issues related to marketing	**Potential for bias** *E.g., creating marketing strategies and content that do not target demographics proportionately* If analytical tools that scan the market and develop marketing approaches to ultimately connect patients with medical products are not trained on representative data, they can limit access to products for potentially already underserved demographics.

[261] https://www.fda.gov/media/165743/download

[262] https://www.fda.gov/media/165743/download

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Feedback and communication tools that facilitate answering questions and gathering input from patients and healthcare providers about medical products *E.g., HCP engagement and experience* Automating responses to HCP questions and providing dynamic feedback *E.g., patient engagement and experience* Streamlining communication with patients and automating follow-up interactions that do not require human interpretation	*E.g., generating communications based on speech or writing patterns that further promote health inequities* Using GenAI to respond to HCP and patient questions or feedback could result in biased or inaccurate responses if not trained on appropriate data based on varying literacy levels, dialects, language spoken, and more, which can perpetuate existing inequities.

Functional component 5: Postmarket monitoring for safety and effectiveness

Oversight and use of medical products in real-world settings to provide care and consistently monitor product safety

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Analytics tools that can provide immediate identification and reporting on efficacy, safety, and compliance *E.g., real-time safety monitoring of medical product use* Analyzing clinical data to identify potential adverse drug reactions or other safety signals from medical products may enable a quick response to protect patient safety.[263] *E.g., automated analysis and identification of patterns in nationwide adverse event reporting* Advanced analytics models on adverse event report data are used to identify potential safety issues for medical products used in clinical settings. *E.g., streamlined pharmacovigilance reporting* Categorizing incidents based on notes, auto-generating feedback insights, and identifying emerging concerns based on data collected from medical product use[264] *E.g., continuous compliance monitoring* Automating compliance audits and ensuring standard operating procedures (SOPs) are followed	**Potential to lower quality of care** *E.g., inaccuracies in postmarket surveillance and monitoring* In devices that operate as clinician augmentation (e.g., screening tools, AI assisting surgical tools), AI could pick up on anomalies, side effects, or adverse reactions in postmarket surveillance and monitoring that are not meaningfully related to the safety of the medical product or fail to identify legitimate anomalies, side effects, or adverse reactions. Similarly, pharmacovigilance analyses that leverage AI may identify "false positives" or "false negatives" as well. Though HCPs and safety monitoring bodies can serve as humans in the loop, there is a potential for overreliance on AI or human error in interpreting AI, which could lead to errors or inaccurate reporting of safety.

There are opportunities to develop and use AI to improve outcomes at each medical product life cycle phase. Every party has an imperative to monitor and mitigate risks alongside innovating. HHS will use the following action plan to safely, responsibly, equitably, and impactfully foster the adoption of AI.

2.6 Action Plan

In light of the evolving AI landscape in medical products and their development, HHS has taken multiple steps across providing regulatory clarity, forming public-private partnerships, and advancing equity in corresponding

[263] https://pmc.ncbi.nlm.nih.gov/articles/PMC9790425/
[264] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112260/

AI technologies to promote responsible AI. The Action Plan below follows the four goals that support HHS's AI strategy: 1. catalyzing health AI innovation and adoption; 2. promoting trustworthy AI development and ethical and responsible use; 3. democratizing AI technologies and resources; and 4. cultivating AI-empowered workforces and organization cultures. For each goal, the Action Plan provides context, an overview of HHS and relevant other federal actions to date, and specific near- and long-term priorities HHS will take. HHS recognizes that this Action Plan will require revisions over time as technologies evolve and is committed to providing structure and flexibility to ensure longstanding impact.

2.6.1 Catalyze Health AI Innovation and Adoption

To help capture the opportunity for AI to transform patient care access and outcomes, HHS plays an active role in furthering innovation and adoption in medical products and across the medical product life cycle. HHS has an opportunity to increase AI uptake in the space by pursuing the following themes of action:

1. Clarifying regulatory oversight of medical products
2. Providing clarity on payment models
3. Fostering public-private partnerships and intergovernmental collaborations to rapidly develop and share knowledge

Below, HHS discusses the context for each area in more detail, corresponding actions to date, and plans to advance AI innovation and adoption across medical products.

1. Clarifying regulatory oversight of medical products:

Context:

There is large growth in the development of AI that can be used across the medical product life cycle. Regarding devices specifically, the rapid growth in the power and availability of new technologies has spurred the development of health information technology applications leveraging AI that fall outside medical device regulations. The 21st Century Cures Act (Cures Act)[265] specifically removed from the FD&C Act[266] the definition of "device" software functions intended for:

- Administrative support of a healthcare facility
- Maintaining or encouraging a healthy lifestyle unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition
- Serve as electronic patient records
- Transferring, storing, converting formats, or displaying test or other device data, results, or findings but not intended to interpret or analyze them
- Certain clinical decision support (CDS) software

The types of CDS software ("non-device CDS") that are not considered devices,[267, 268] such as applications which support or provide recommendations to an HCP and:

- Do not acquire, process, or analyze medical images, signals, or patterns
- Do not display, analyze, or print medical information beyond what would normally be communicated between healthcare professionals
- Do not provide a specific output or directive

[265] https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act

[266] https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act

[267] https://www.fda.gov/medical-devices/software-medical-device-samd/your-clinical-decision-support-software-it-medical-device

[268] Some CDS software may still be regulated as devices if they meet the definition of "device" in the FD&C Act. 21 USC 321(h). Any software or AI *intended* to diagnose, cure, mitigate, treat, or prevent disease is a device.

- Do not require the healthcare professional to rely primarily on the recommendations by providing the basis of the recommendations to inform decision-making

ASTP's HTI-1 Final Rule addresses the availability of AI in certain certified EHR systems, which, as of 2021, have been adopted by 96% of hospitals and 78% of physician offices across the country.[269] The HTI-1 Final Rule does not create an approval process per se, but it does establish policies that require transparency on the part of certain certified health IT products regarding the technology offered in such products. Starting on January 1, 2025, regulations finalized in the final rule require the availability of specific "source attribute" information for any decision support intervention technologies certified to 45 CFR 170.315(b)(11) (including AI-based decision support interventions) offered as part of the health IT product. These requirements apply to AI-based technologies regardless of device definitions, use cases (e.g., clinical versus administrative), or risk categories. As the growth of AI in health IT (e.g., EHRs) continues, there will be a need for greater clarity on regulatory boundaries and applicability to minimize business uncertainty that may hinder innovation.

While this theme of action may be more pertinent to devices than drugs and biological products, the wide availability of AI is spurring growth across all medical products. As developers make investment and product roadmap decisions, there is a growing need for further clarity on the definitions that determine regulatory pathways that could affect the cost and timing of device, drug, and biological product development.

HHS actions to date (non-exhaustive):

- **FDA's Guidance on Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations**[270] provides recommendations regarding the contents of marketing submissions for devices that include AI-enabled device software functions including documentation and information that will support FDA's evaluation of safety and effectiveness. The recommendations reflect a comprehensive approach to the management of risk throughout the device total product life cycle (TPLC). To support the development of appropriate documentation for FDA's assessment of the device, this draft guidance also proposes recommendations for the design, development, and implementation of AI-enabled devices that manufacturers may wish to consider using throughout the TPLC.
- **FDA's Guidance on Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products**[271] provides recommendations to sponsors and other interested parties on the use of AI to produce information or data intended to support regulatory decision-making regarding safety, effectiveness, or quality for drugs. Specifically, this guidance provides a risk-based credibility assessment framework that may be used for establishing and evaluating the credibility of an AI model for a particular context of use (COU).
- **FDA's Guidance on Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions**[272] provides recommendations for predetermined change control plans (PCCPs) tailored to AI-enabled devices and intends to support iterative improvement through modification to AI-enabled devices while ensuring safety and effectiveness.

[269] https://www.healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records

[270] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/artificial-intelligence-enabled-device-software-functions-lifecycle-management-and-marketing

[271] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological

[272] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence

- **FDA's CDS Software Guidance for Industry and FDA Staff**[273] provides clarification on the 21st Century Cures Act legislation that excludes certain CDS software from the FDA's device jurisdiction. This helps elucidate the complexities of certain unregulated uses of AI in healthcare technology.
- **FDA's "Artificial Intelligence and Medical Products: How CBER, CDER, CDRH, and OCP Are Working Together" paper**[274] specifies how the Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), Center for Devices and Radiological Health (CDRH), and Office of Combination Products are working together to identify steps to foster collaboration, develop regulations, promote best practices, and support corresponding research efforts.
- **FDA's Digital Health and Artificial Intelligence Glossary—Educational Resource**[275] is a publicly available resource that defines common terms in digital health, AI, and ML to provide internal and external consistency and education.
- **AHRQ's Clinical Decision Support Innovation Collaborative** has been advancing patient-centered clinical decision support (PC CDS), including exploring the impacts of AI on PCCDS and conducting pilot projects.[276]

HHS near-term priorities:

- Continue to issue guidelines, supporting materials (e.g., FAQs), and/or discussion papers regarding the use of AI in medical product development and in medical products to provide further recommendations.
- Consider new resourcing opportunities to research AI and CDS, including ways to understand better the benefits and risks of using clinical data in CDS software.

2. Providing clarity on payment models:

Context:

Across clinical disciplines (e.g., radiology and pathology), some devices incorporate AI with proven effectiveness; however, because many devices do not have established payment, full uptake potential has yet to be realized.[277] Healthcare delivery payment and coverage policies can influence the economics underlying the adoption of AI. While some medical products may have clear efficiency or productivity return on investment benefits where there is a market, there can be disconnects between patient benefits and financial incentives in the complex way healthcare gets paid for in the U.S. Purchasers and payers need evidence with outcomes and/or endpoints for patient populations relevant to coverage decisions and indications for use relevant to payers and beneficial for commercialization and patient access. Without a clear path for uptake in clinical settings, medical device developers may be less incentivized to continue innovating on these types of products.

In general, for an item or service to be considered for Medicare coverage, the item or service must fall within at least one benefit category established in the Social Security Act (the Act), the item or service must not be specifically excluded by the Act, and the item or service must be "reasonable and necessary for the diagnosis or treatment of illness or injury."[278] CMS may issue a National Coverage Decision (NCD) to describe the nationwide conditions for Medicare coverage for a specific item or service. Without an NCD, items and services are covered on a claim-by-claim basis at the discretion of the Medicare Administrative Contractors (MACs) or through a Local Coverage Determination. As of May 2024, CMS has established payment for at least eight AI-enabled devices through Current Procedural Terminology (CPT®) and New Technology Add-

[273] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software
[274] https://www.fda.gov/media/177030/download?attachment
[275] https://www.fda.gov/science-research/artificial-intelligence-and-medical-products/fda-digital-health-and-artificial-intelligence-glossary-educational-resource
[276] https://cdsic.ahrq.gov
[277] https://www.massbio.org/wp-content/uploads/2024/09/FINAL-Vision-2030-Strategy-Report.pdf
[278] https:/www.cms.gov/medicare/coverage/councilontechinnov/downloads/innovators-guide-master-7-23-15.pdf

On Payment (NTAP) under the Medicare Inpatient Prospective Payment System (IPPS),[279] less than 5% of FDA-authorized AI-based products.[280] CMS also established payment pathways for hospital outpatient departments through separate payment of software-as-a-service add-on codes in 2022. Over time, the growth of value-based purchasing payment models may provide more built-in financial incentives for investment in AI in healthcare, but the growth of such programs is not rapid. Further clarifying existing pathways could spur established payment for more AI-enabled devices.

HHS actions to date (non-exhaustive):

- **CMS's NTAP**[281] provides for an add-on payment for certain new devices under the Medicare Inpatient Prospective Payment System (IPPS),[282] including those leveraging AI, with a few examples dating back to 2020 (e.g., ContactCT by Viz.ai, AI-driven triage software for large-vessel occlusion). Since then, additional AI software developers (e.g., RapidAI, AIdoc, Avicenna) have also been granted NTAP status.
- **CMS' Transitional Coverage for Emerging Technologies (TCET) pathway** helps people with Medicare access the latest medical advances, enables doctors and other clinicians to provide the best care for their patients, and benefits manufacturers who create innovative technologies.[283]
- **CMS' Medicare Pharmaceutical and Technology Ombudsman** has been in place since late 2017. This ombudsman receives and assists with inquiries and complaints from pharmaceutical, biotechnology, medical device, diagnostic product manufacturers, and other stakeholders regarding coverage, coding, and/or payment for products covered by Medicare or for which Medicare coverage is being sought.[284]

HHS near-term priorities:

- Convene HHS divisions (e.g., CMS, NIH, FDA, and ASTP) to align on benefits, risks, and potential definitions of standardized, future-proof payment pathways for AI-enabled medical devices.
- Expand the **Early Payer Feedback Program** to shorten the time to payment and coverage determinations with commercial and government insurers.
- Issue guidelines to healthcare payers, providers, and other stakeholders on the pathways available to establish payment for AI-enabled devices.

HHS long-term priorities:

- Develop clear payment pathways for AI-enabled medical devices in the public sector to potentially spur similar activity in the private sector.
- Iteratively reevaluate guidelines and payment pathways for AI-enabled medical devices as healthcare technology transforms to continue fostering adoption while mitigating risks.

3. **Fostering public-private partnerships and intergovernmental collaborations to rapidly develop and share knowledge:**

Context:

Regulatory bodies worldwide are taking different approaches to publish guidelines regarding AI in medical products. Medical product developers, manufacturers, and distributors who aim to serve patients globally could pursue innovation more efficiently with cooperative standards and guardrails to follow. Regulatory processes that ensure the safety and effectiveness of medical products are critical to safeguarding the American public, and FDA's medical product centers intend to continue administering programs that accelerate

[279] https://www.nature.com/articles/s41746-022-00609-6/tables/1

[280] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

[281] https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/new-medical-services-and-new-technologies

[282] https://www.cms.gov/cms-guide-medical-technology-companies-and-other-interested-parties/payment/ipps

[283] https://www.cms.gov/newsroom/fact-sheets/final-notice-transitional-coverage-emerging-technologies-cms-3421-fn

[284] 42 U.S.C. Section 1395b-9, https://www.cms.gov/center/special-topic/ombudsman/medicare-pharmaceutical-and-technology-ombudsman

innovation and provide regulatory guidelines for the use of AI. Furthermore, by continuing and building upon its interaction directly with the private sector, HHS can share knowledge in a way that unlocks further advancements in AI in medical products and across the medical product life cycle.

HHS actions to date:

- **FDA's engagement in public-private partnerships (PPPs),**[285, 286, 287] through collaborations with other government, academic, scientific, patient, and private sector organizations, advances science and innovation in how medical products are developed, evaluated, and manufactured. These ongoing efforts encourage the development of new tools, including AI, to facilitate innovation across the medical product life cycle. Example PPPs that include potential AI-specific focus areas are:
    - **BioFabUSA**[288] works to integrate innovative cell and tissue cultures with advances in biofabrication, automation, robotics, and analytical technologies to create disruptive research and development tools and FDA-compliant volume manufacturing processes.
    - **The National Institute for Innovation in Biopharmaceuticals (NIIMBL)**[289] facilitates innovative manufacturing technologies and workforce development programs to foster efficiencies and impact in the life sciences industry.
    - **Critical Institute Path (C-Path)**[290] is a non-profit organization dedicated to improving and streamlining drug development through fostering collaboration between private sector industry executives and scientists, academic researchers, regulators, and patient groups.
    - **Clinical Trials Transformation Initiative (CTTI)**[291] brings together organizations and individuals representing academia, clinical investigators, government and regulatory agencies, private sector industry, IRBs, patient advocacy groups, and others to develop evidence-based solutions to clinical research challenges.
- **NIH's Advancing Health Research through Ethical, Multimodal Artificial Intelligence (AI) Initiative**[292] funds the development of ethically focused and data-driven multimodal AI approaches to more closely interpret and predict complex biological and behavioral systems and model intricate health systems to enhance our understanding of health and the ability to detect and treat human diseases.
- **FDA's Artificial Intelligence Program—Research on AI-based medical devices**[293] relies on the CDRH conducting regulatory science research to ensure patient access to safe and effective medical devices using AI. Specific focus areas include methods to enhance model training, minimize bias, and develop methods to track safety postmarket.
- FDA, NIH, and NSF launched the **Foundations for Digital Twins as Catalyzers of Biomedical Technological Innovation (FDT-BioTech) program** to catalyze biomedical innovation through synthetic data, which facilitates clinical trials by providing control data that may be challenging to obtain through traditional participant recruitment.[294]
- **Across NIH, its institutes, centers, and offices are funding research**[295] **to apply AI** in many disease contexts including in wearable technology to help monitor and screen cognitive impairment,[296] to detect neurological disease through retinal imaging, and identify patients with potential substance misuse disorders.

---

285 https://www.fda.gov/emergency-preparedness-and-response/innovative-technologies/public-private-partnerships

286 https://www.fda.gov/drugs/science-and-research-drugs/scientific-public-private-partnerships-and-consortia

287 https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-research-and-partnerships

288 https://www.fda.gov/emergency-preparedness-and-response/innovative-technologies/public-private-partnerships

289 https://www.fda.gov/emergency-preparedness-and-response/innovative-technologies/public-private-partnerships

290 https://c-path.org/c-path-awarded-fda-grant-to-establish-public-private-partnership-to-advance-treatments-for-rare-neurodegenerative-diseases/

291 https://www.fda.gov/patients/learn-about-fda-patient-engagement/fda-patient-engagement-partnerships

292 https://datascience.nih.gov/sites/default/files/MAI-Solicitation-outline.pdf

293 https://www.fda.gov/medical-devices/medical-device-regulatory-science-research-programs-conducted-osel/artificial-intelligence-program-research-aiml-based-medical-devices

294 https://new.nsf.gov/funding/opportunities/fdt-biotech-foundations-digital-twins-catalyzers-biomedical

295 https://grants.nih.gov/funding/find-a-fit-for-your-research-nih-institutes-centers-offices

296 https://www.nia.nih.gov/research/milestones/diagnosis-assessment-and-disease-monitoring/enabling-tech-scalable-wearables

- **NIH's National Cancer Institute (NCI)-DOE collaboration as a part of the Cancer Moonshot**[297] accelerates advances in precision oncology and scientific computing, including the use of AI.

HHS near-term priorities:

- Leverage and continue to build upon existing initiatives around the use of AI in medical products and across the medical product life cycle.
- Explore approaches to a PPP that advances innovation, commercialization, and risk-mitigation methods for AI in medical products and across the medical product life cycle to help promote safe, responsible, fair, privacy-protecting, and trustworthy AI in the space as articulated in E.O. 11410.[298]
- Evaluate approaches to continue expanding the **Total Product Life Cycle Advisory Pilot (TAP)**[299] and **Early Payer Feedback Program (EPFP)** to accelerate the identification of innovation, adoption, and commercialization barriers to AI, especially for developers less familiar with device marketing authorization processes and payer coverage decision-making.
    - Coordinate with strategic investments targeting underinvested TAs.

HHS long-term priorities:

- Continue monitoring and evaluating trends and emerging issues to detect potential knowledge gaps and opportunities that may permit timely adaptations that provide clarity for using AI in the medical product life cycle.
- Continue working closely with global collaborators to promote international cooperation on standards, guidelines, and best practices to encourage collaboration in using and evaluating AI across the medical product landscape.
- Explore resourcing for developing educational initiatives to support regulatory bodies, healthcare professionals, patients, researchers, and private sector industry as they navigate the safe and responsible use of AI in medical products and their development.
- Explore resourcing to support regulatory science efforts to develop additional methodologies for evaluating AI algorithms, identifying and mitigating bias, and ensuring their robustness and resilience to changing clinical inputs and conditions.

2.6.2 Promote Trustworthy AI Development and Ethical and Responsible Use

HHS will promote the trustworthy, ethical, and responsible use of AI in medical products or across the medical product life cycle as follows:

1. Refining regulatory frameworks to address adaptive AI technologies in medical devices
2. Promoting equity in AI deployment to bolster safe and responsible use
3. Addressing AI-enabled software outside current device regulatory authorities
4. Fostering private or public mechanisms for quality assurance of health AI

Below, HHS discusses the context of each area in more detail, corresponding actions to date, and forward-looking plans to ensure AI use is trustworthy and safe for use in medical products and across the medical product life cycle.

[297] https://datascience.cancer.gov/collaborations/nci-department-energy-collaboration

[298] https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

[299] https://www.fda.gov/medical-devices/how-study-and-market-your-device/total-product-life-cycle-advisory-program-tap

1. Refining regulatory frameworks to address adaptive AI in medical devices

Context:

The FDA's traditional paradigm of medical device regulation may not have been designed for adaptive AI technologies that could continuously change and optimize device performance in real time to improve patient healthcare. The current regulatory approach is to monitor the performance and safety of a device as configured at marketing authorization[300] and may not address adaptive technologies such as AI, which may deviate considerably from what was originally presented for authorization. Most FDA-authorized medical devices come through the 510(k)-pathway based on demonstrating substantial equivalence to a lawfully marketed "predicate" device. As the complexity of such technologies increases, more specific and explicit premarket demonstrations of the safety and effectiveness of such products may help account for adaptive AI and other technologies. The highly iterative, autonomous, and adaptive nature of these tools may benefit from a total product life cycle (TPLC),[301] a regulatory approach that facilitates a rapid product improvement cycle and allows these devices to improve while continually providing effective safeguards. With appropriately tailored regulatory oversight, AI can deliver safe and effective functionality that improves the quality of patient care.

HHS actions to date (non-exhaustive):

- **FDA's Action Plan for Artificial Intelligence and Machine Learning Based Software as a Medical Device (SaMD)**[302] from 2021 outlined a multipronged approach to advance the agency's oversight of these technologies. FDA has:
    - Issued draft guidance on marketing submission recommendations for predetermined change control plans for AI-enabled device software functions.[303]
    - Published Guiding Principles on Good Machine Learning Practice for Medical Device Development[304] with our partners from Health Canada and the U.K.'s Medicines and Healthcare products Regulatory Agency (MHRA).
    - Hosted a public workshop on Transparency of AI-enabled Medical Devices.[305]
    - Released a "Spotlight: Digital Health Regulatory Science Opportunities." The Spotlight highlights common digital health interest areas, including AI and ML, among other topics. It presents these current regulatory science areas of interest in digital health for all to consider.[306]
- **ARPA-H's Performance and Reliability Evaluation for Continuous Modifications and Useability of Artificial Intelligence (PRECISE-AI) program**[307] funds investigation to develop technology that can detect when AI-enabled tools used in clinical care settings are out of alignment with underlying training data and auto-correct them.

HHS near-term priorities:

- Explore policies for using AI to produce information for regulatory decision-making, including potential approaches to defining questions of interest, contexts of use, model risks, and model output credibility.
- Explore "model card" approaches across various regulatory frameworks for AI.

---

300 https://www.fda.gov/medical-devices/510k-clearances/medical-device-safety-and-510k-clearance-process

301 https://www.fda.gov/about-fda/cdrh-transparency/total-product-life-cycle-medical-devices The use of AI in the medical product life cycle for the development of drugs, biological products, devices, or combination products may differ. For example, for drugs and biological products, the end product is typically the drug or biological product itself, which will generally not include AI in that end product. For devices, the end product is the device, which may itself be AI-enabled. When describing the life cycle of a medical device, including AI-enabled devices, the term "Total Product Life Cycle," or TPLC, is often used. For more information, see Total Product Life Cycle for Medical Devices, September 6, 2023 (link at the beginning of this footnote).

302 https://www.fda.gov/media/145022/download

303 https://www.fda.gov/regulatory-information/search-fda-guidance-documents/predetermined-change-control-plans-medical-devices

304 https://www.fda.gov/media/153486/download

305 https://www.nature.com/articles/s41746-023-00992-8

306 https://www.fda.gov/media/162644/download

307 https://arpa-h.gov/research-and-funding/programs/precise-ai

- Develop standards, guidelines, and innovative science-based approaches to assess the safety, effectiveness, and/or performance of AI-enabled medical devices.
- Explore resourcing for research on evaluating and monitoring AI performance in medical devices.
- Explore resourcing for evaluating and using robust AI tools to model drift in medical devices as a potential complement to the ARPA-H PRECISE-AI program.
- Incorporate AI for regulatory submissions by sponsors and FDA internal review processes.

HHS long-term priorities:

- Continue refining and developing considerations for evaluating the safe, effective, responsible, and ethical use of AI in the medical product life cycle (e.g., AI provides adequate transparency and addresses safety, effectiveness, and cybersecurity concerns).

2. Promoting equity in AI deployment to bolster safe and responsible use

Context:

FDA is taking steps to advance health equity in the context of medical products.[308] ASTP requirements on certified health IT products do include health equity components;[309] However, the scope of ASTP regulations is limited to certified health IT or products, including certified health IT. As the use of AI in medical products and across the medical product life cycle continues to increase, HHS can consider approaches to bolster health equity in this area.

HHS actions to date:

- **FDA's Artificial Intelligence and Medical Products: How CBER, CDER, CDRH, and OCP Are Working Together paper**[310] discusses how FDA's medical product centers work closely with developers, patient groups, academia, global regulators, and other stakeholders to cultivate a patient-centered regulatory approach emphasizing collaboration and health equity. The paper also describes FDA's support for projects considering health inequities associated with using AI in medical product development to promote equity and ensure data representativeness, leveraging ongoing diversity, equity, and inclusion efforts.
- **ASTP's blog post Embracing Health Equity by Design**[311] discusses a multifaceted approach to equity in healthcare IT. It includes using the right data, selecting the appropriate tools, and ensuring interoperability between systems to reduce bias and ensure all groups are represented proportionately in health technology.
- **AHRQ's Digital Healthcare Equity Framework and Practical Guide for Implementation** helps organizations intentionally consider equity in developing and using digital healthcare technologies and solutions. The Guide is a resource for digital healthcare developers, vendors, healthcare systems, clinical providers, and payers. It includes steps and real-world examples for advancing equity across the Digital Healthcare Life Cycle phases.[312]

Applicable federal laws to date:

- **Section 1557 of the Affordable Care Act** prohibits discrimination based on race, color, national origin, sex, age, and disability in certain health programs and activities through patient care decision support tools, including AI.[313] (*See Appendix B for additional, non-exhaustive federal policies and regulations*)

[308] www.fda.gov/media/180608/download?attachment
[309] https://www.healthit.gov/buzz-blog/health-it/embracing-health-equity-by-design
[310] https://www.fda.gov/media/177030/download?attachment
[311] https://www.healthit.gov/buzz-blog/health-it/embracing-health-equity-by-design
[312] https://digital.ahrq.gov/health-it-tools-and-resources/digital-healthcare-equity
[313] https://www.hhs.gov/civil-rights/for-individuals/section-1557/index.html

HHS near-term priorities:

- Explore resourcing for internal and external projects, highlighting different points where bias can be introduced in the AI development life cycle and how it can be addressed through risk management.
- Disseminate research on best practices for documenting and ensuring that data used to train and test AI models are fit for use and adequately represent the target population to help bolster equity considerations that promote safe and responsible AI use.
- Explore resourcing for projects considering health inequities associated with using AI in medical product development to promote equity and ensure data representativeness, leveraging ongoing diversity, equity, and inclusion efforts, to help ensure ethical and trustworthy use of AI in medical products and their development.
- Explore resourcing for clinical trials leveraging AI to address areas of unmet need or those where the pipeline does not meet the burden.

HHS long-term priorities:

- Continue to explore resourcing for internal and external projects, highlighting different points where bias can be introduced in the AI development life cycle and how it can be addressed through risk management.

3. Addressing AI-enabled software outside current device regulatory authorities

Context:

An increasing number of AI tools in health IT could fall outside FDA regulation, including certain EHR-integrated AI decision support tools (e.g., appointment no-show prediction) and AI algorithms deployed by health plans and insurance issuers for utilization management and prior authorization. Authority over the regulation of health IT, which is not medical devices, belongs partly to the ASTP/ONC. Tools that do not meet the FDA's device definition may not undergo regulatory review, validation, or testing.[314] This is an area that HHS will continue to monitor closely.

HHS actions to date:

- **ASTP/ONC's HTI-1 Final Rule**[315] finalized policies that require certain certified health IT (such as EHR health IT products certified to the certification criterion at 45 CFR 170.315(b)(11)) to enable users to access information about the design, development, training, and evaluation of AI (called predictive decision support interventions or PDSIs) to help users determine whether the tool is appropriate for their care setting and patient population.
- **FDA CDS Software Guidance for Industry and FDA Staff**[316] provides clarification on the 21st Century Cures Act legislation that excludes certain CDS software from the FDA's device jurisdiction, which helps elucidate the complexities of certain uses of AI in healthcare technology that are not regulated as devices.

HHS near-term priorities:

- Assess mechanisms to ensure appropriate oversight of AI outside FDA regulatory authority and continuously monitor advances in the ecosystem.
- Explore approaches for:
    - "Model card" information for AI-based technologies outside of FDA's jurisdiction
    - Bolstering the validation of AI-based models with clinical data
    - Including health equity considerations in regulatory pathways

[314] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software

[315] https://www.healthit.gov/topic/laws-regulation-and-policy/health-data-technology-and-interoperability-certification-program.

[316] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software

- Public-private collaboration models for rigorous, standards-based, pre-, and postmarket quality assurance of AI-based technologies outside of FDA's jurisdiction

HHS long-term priorities:

- Iteratively monitor and reevaluate regulatory oversight mechanisms of AI in medical and health technologies outside of FDA's jurisdiction as the field rapidly evolves.
- Explore opportunities to collect feedback about AI in medical and health technologies outside FDA's jurisdiction to monitor the potential impacts of such technologies on healthcare.

4. Fostering private or public mechanisms for quality assurance of health AI

Context:

Despite the promise of AI tools in medicine, the ability to prospectively test AI tools across diverse datasets and deploy AI in multiple clinical care settings to ensure consistency, accuracy, and generalizability in improving health outcomes can be limited by the availability of such datasets and inconsistent monitoring in clinical use. Testing of AI to identify potential biases, disparities, or inconsistencies in AI model performance and optimizing AI models for diverse healthcare environments can be improved through increased availability of data and improved monitoring capabilities. The absence of standardized quality assurance (QA) protocols designed to evaluate performance in real-world settings to ensure continued patient and provider safety increases the risk of inconsistent implementation across sites and unintended consequences.[317, 318] Even AI tools that received regulatory clearance for clinical use may underperform when deployed in new clinical settings due to poor generalization or when used for a purpose other than its authorized intended use. These cases highlight the challenges AI tools face in medicine due to biases in development data (e.g., training, tuning, internal test sets used by the developer to create the tool) and the potential distribution shifts in the characteristics of external, previously unused test sets or patient cases. For the safe and effective integration of AI tools into the clinical workflow, "transparency[319] from manufacturers about the development process," and the implementation of QA programs could be necessary.[320]

HHS actions to date:

- **FDA's collaboration with the Department of Veterans Affairs,**[321] announced in October 2024, will be an "interagency testing ground" for healthcare-related AI tools. The lab will "serve as an asset for federal agencies and the private sector 'to be able to test applications of AI in a virtual lab environment to ensure not only that they work and that they're safe and effective for veterans and patients,' but that they also 'adhere to trustworthy AI principles,'" according to VA Undersecretary for Health Shereef Elnahal.

HHS near-term priorities:

- Collaborate with public and private networks on testing health AI to provide shared resources and infrastructure that encourage safe and effective development, transparency, reporting, and ongoing monitoring of health AI.
- Consider supporting guidelines and educational tools to help AI developers as they work toward safety, security, and trust while creating AI technologies for use in medical products and across the medical product life cycle.

---

317 https://pmc.ncbi.nlm.nih.gov/articles/PMC5438240/

318 https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.16188

319 See FDA's "Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles" for more information on "transparency" in this context: https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices-guiding-principles

320 https://pmc.ncbi.nlm.nih.gov/articles/PMC10928809/#ubae003-B19

321 https://www.nextgov.com/artificial-intelligence/2024/10/va-announces-creation-new-ai-testing-ground-fda/400681/?oref=ng-homepage-river

HHS long-term priorities:

- Explore resourcing to develop regulatory science approaches to assess the accuracy and reliability of AI models once deployed in a healthcare environment.

2.6.3 Democratize AI Technologies and Resources

To effectively capture the value of AI in medical products across the medical product life cycle while mitigating associated risks, technology uptake and innovation could benefit from equitable access throughout the ecosystem across a diverse set of players (e.g., medical technology companies, academia, non-profits, and public sector entities) and stakeholders (e.g., from different demographic backgrounds). Without such accessibility, capturing the full value potential of AI in the space might not be feasible or fully account for risks. HHS plans to play a key role in mitigating this by integrating equity principles into the expansion of AI in medical products along the following key themes of action:

1. Enabling collaborative development through public engagement
2. Aligning standards and information-sharing mechanisms across research and healthcare delivery

Below, HHS discusses the context of each theme of action in more detail, corresponding actions to date, and plans to ensure equitable access to AI technologies and resources in medical products.

1. Enabling collaborative development through public engagement

Context:

Increased stakeholder collaboration could democratize AI technologies and best practices in medical products and across the medical product life cycle. A lack of collaboration between stakeholders (e.g., private sector industry, STLTs, academia, and the general public) and intentional public engagement throughout the medical products life cycle could limit the potential of AI to be equitably adopted broadly across medical products and their development.

HHS actions to date:

- **NIH's AIM-AHEAD Program** is designed to support mutually beneficial triadic partnerships among (1) local, state, and tribal accredited health departments; (2) limited-resource higher education institutions; and (3) a data-science-oriented organization with an accessible data library to collaboratively conduct health-equity-related AI studies.[322] These critical and trusted organizations can benefit from enhancing their AI capabilities to advance public health, from early detection and monitoring, predictive analytics, disease surveillance and monitoring, and outbreak detection to healthcare resource allocation and personalized interventions. Partnerships among public health department professionals, academic researchers, and data-science/AI experts can further leverage data-driven insights that contribute to more effective and efficient public health strategies to improve community health outcomes.
- **The Department of Energy and the NIH's collaboration through the National Artificial Intelligence Research Resource (NAIRR) Secure Pilot** will "enable research that involves sensitive data, which require special handling and protections. The NAIRR Secure pilot will assemble exemplar privacy/security-preserving resources (e.g., data enclaves, secure compute resources, and privacy-preserving tools) and develop requirements for the future NAIRR Secure."[323]

HHS near-term priorities:

- Develop a vision and framework for incorporating public voices in all parts of the medical products life cycle.[324]
- Convene a public-private community of practice for sharing best practices and identifying enablers/barriers to AI adoption in clinical studies.
- Refine and develop a more robust STLT engagement strategy regarding medical products where appropriate to ensure best practices on AI are shared between all levels of government.

HHS long-term priorities:

- Offer secure sandboxes[325] to encourage collaborative innovation in developing and using AI for medical products.
- Engage in public and private collaborations, fostering long-term relationships between the private sector industry, providers, and the public that can be tapped for co-creation.
- Explore resourcing for multi-institutional collaboration mechanisms, especially those potentially under-resourced organizations that could benefit from knowledge or infrastructure sharing.

2. Aligning standards and information-sharing mechanisms across research and healthcare delivery

Context:

Clear standards for data, metadata, and pathways to share information can make AI innovation easier to access. A lack of clear standards can make data across private sector industries, academia, non-profits, governments, and other players unusable or non-transferable to AI models, stifling AI uptake in medical products and across the medical product life cycle.[326] Barriers to sharing data can be more prohibitive to innovation for stakeholders with less access to resources than for those with higher resources who can fund data collection or data cleaning activities.

[322] https://datascience.nih.gov/artificial-intelligence/aim-ahead

[323] https://nairrpilot.org/nairr-secure

[324] https://osp.od.nih.gov/policies/novel-and-exceptional-technology-and-research-advisory-committee-nextrac. This is the current charge of an NIH FACA called the NExTRAC.

[325] See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan.

[326] https://pmc.ncbi.nlm.nih.gov/articles/PMC2213488/

HHS actions to date (non-exhaustive):

- **ARPA-H's Imaging Data Partnership** with the CDRH of FDA aims to streamline access to affordable, high-quality medical imaging data.[327] The agencies work together to develop a medical imaging data marketplace to accelerate AI and ML innovation by removing barriers to obtaining data that align with regulatory quality standards and appropriately represent the relevant portions of the U.S. population.
- **ARPA-H's Biomedical Data Fabric toolbox** seeks to facilitate the connection of biomedical research data from thousands of sources, advancing the collection and usability of biomedical datasets originating from thousands of different research labs, clinical care centers, and other data sources and accelerating technical innovation across the health ecosystem.[328] By (1) lowering barriers to high-fidelity, timely data collection in computer-readable forms, (2) preparing for multisource data analysis at scale, (3) advancing intuitive data exploration, (4) improving stakeholder access while maintaining privacy and security measures, and (5) ensuring generalizability of biomedical data fabric tools across disease types, ARPA-H is democratizing access to data. These data must be findable, accessible, interoperable, and reusable. NIH's Generalist Repository Ecosystem Initiative (GREI) supports seven established generalist repositories that work together to establish consistent metadata, develop use cases for data sharing and reuse, and train and educate researchers on how to share and reuse data, including for the development and use of AI.[329]
- **NIH's Toward an Ethical Framework for Artificial Intelligence in Biomedical and Behavioral Research: Transparency for Data and Model Reuse Workshop** focused on highlighting the importance of standardizing the safe shareability of synthetic data, data sharing for general reuse, and multimodal data, which can lead to transformational product development if leveraged in AI tools.[330]

HHS near-term priorities:

- Release draft guidelines on data-sharing principles consistent with the **HHS Data Strategy,** including common approaches to structuring data and metadata and clarity around what data types can be published and shared.[331]
- Offer secure sandboxes[332] to spur collaborations in data sharing and standards development.
- Develop open-industry standards and open-source tooling and infrastructure for registries to leverage AI to support device pre- and postmarket submission requirements, cross-standard data mapping, and de-identification to develop AI-ready datasets and tooling.
- Accelerate work with standards development organizations and industry collaborations on standards to support AI development and use across the life cycle.
- Accelerate alignment of federally funded research data standards (semantic, format, and transport) with HHS-adopted standards for EHRs, healthcare providers, and payers (e.g., USCDI, USCDI+, HL7, FHIR, and CARIN).

HHS long-term priorities:

- As the landscape changes for public access to research results, data management, and sharing, HHS may need to build added capacity to assist key players in refining standards for both.

[327] https://arpa-h.gov/news-and-events/arpa-h-announces-medical-imaging-data-partnership-fda

[328] https://arpa-h.gov/research-and-funding/programs/arpa-h-bdf-toolbox

[329] https://datascience.nih.gov/data-ecosystem/generalist-repository-ecosystem-initiative

[330] https://datascience.nih.gov/sites/default/files/ai-meetings/NIH-Transparency-Workshop-Report-v6-FINAL-updated-09-16-24-508.pdf

[331] https://cdo.hhs.gov/s/hhs-data-strategy

[332] See Appendix A: "Glossary of terms" for the definition of "sandbox" used in this Plan

2.6.4 Cultivate AI-Empowered Workforces and Organization Cultures

Without a sufficient supply of talent in AI to enable innovation at scale in medical products and across the medical product life cycle, widescale adoption and effective uptake may not be feasible. To that end, HHS plans to spur workforce development externally and internally to empower continued responsible, safe innovation of AI across the medical product life cycle by focusing on key themes of actions:

1. Improving training in the governance and management of AI in medical products
2. Developing and retaining AI talent related to medical products

Below, HHS discusses the context of this goal in more detail, corresponding actions to date, and plans to cultivate AI-empowered workforces and organizational cultures in medical products.

1. Improving training in the governance and management of AI in medical products

Context:

Most individuals involved in AI will be responsible for managing and using such technologies rather than developing them. Ensuring that the medical product ecosystem (including developers, clinicians, and patients) gets the most out of AI will require focusing not just on the technologies themselves but also on their implementation, workflow integration, and life cycle management. Training to enable the research workforce to responsibly manage and use such technologies will be critical to harnessing AI to advance medical products.

HHS actions to date (non-exhaustive):

- **FDA's blog entry, "A Lifecycle Management Approach Toward Delivering Safe, Effective AI-Enabled Health Care,"**[333] provides an overview of one potential approach to developing, validating, and managing ongoing governance of AI use in medical devices to maintain their safety and effectiveness. This approach could provide a foundation for HHS to build upon to develop further best practices for training on governance and management of AI in medical devices and during their development.

HHS near-term priorities:

- Explore targeting resources, training, and workshops to include governance and management of AI technologies in clinical research, including in clinical trial design and management.

HHS long-term priorities:

- Develop internal data science, computer science, and AI talent related to medical products through targeted internal trainings or apprenticeship programs.

2. Developing and retaining AI talent related to medical products

Context:

To harness the potential of AI, the private sector industry, government, academia, non-profits, and other involved parties may need a strong pipeline for a diverse workforce capable of developing and embedding AI to enhance medical products and their development. Professionals from all backgrounds will need baseline knowledge to develop and apply AI safely, responsibly, and effectively. Therefore, developing and retaining AI talent related to medical products could be critical to growing and maintaining innovation.

[333] https://www.fda.gov/medical-devices/digital-health-center-excellence/blog-lifecycle-management-approach-toward-delivering-safe-effective-ai-enabled-health-care

HHS actions to date:

- **FDA's STEM Outreach, Education, and Engagement Program** seeks to provide educational opportunities to prospective scientists, raise awareness of the FDA as a science-based agency, expose students to the broad scope of regulatory science and its impact on our lives, inspire future innovators to pursue the wide range of scientific careers that make up the field of regulatory science at the FDA, and recruit and hire scientists.[334] Though the program is generally oriented toward the FDA, it enhances the overall talent ecosystem and can explore additional focuses related to AI.
- **NIH's Bridge2AI program** creates flagship datasets based on ethical principles, associated standards and tools, and skills and workforce development to address grand challenges in biomedical and behavioral research that require AI analysis.[335]
- **FDA's scientific internships and fellowships** offer undergraduate and graduate students the chance to explore careers related to research, regulatory science, and other STEM fields that develop potential future FDA and other technical talent in the workforce.[336] Though the program is generally oriented toward the FDA, it enhances the overall talent ecosystem and can help promote the exploration of additional focuses related to AI across medical products.
- **HHS integrated AI into enterprise activities** (see the Internal Operations chapter) and released a public tracker of all use cases.[337] As of 2023, there were 164 AI use cases across HHS and its divisions, including deduplicating data, detecting adverse events, monitoring safety, managing signal detection, visualizing data, and analyzing texts. For example, the FDA is exploring the use of AI in various fields, including deduplicating non-public adverse event data in the FAERS and identifying novel terms for opioid-related drugs using the Term Identification and Novel Synthetic Opioid Detection and Evaluation Analytics tool, which uses publicly available social media and forensic chemistry data to identify novel referents to drug products in social media texts.[338]

HHS near-term priorities:

- Expand internship and apprenticeship programs to incorporate AI-specific roles related to medical products and their development.
- Explore additional resourcing for existing outreach, education, and engagement programs to incorporate AI-specific content, particularly those related to medical products and their development.
- Evaluate the expansion of **NIH's AIM-AHEAD Program** to include recruitment and training for AI expertise in clinical research.

2.7 Conclusion

AI can be a medical device, be part of a medical device, enhance the design and conduct of clinical trials, streamline manufacturing and supply chains, and bolster postmarket surveillance and monitoring of medical products, ultimately improving patient care and accessibility to innovative medical products. However, the rapid advancement of AI also presents challenges that should be addressed. HHS's balanced approach aims to foster AI innovation while maintaining robust regulatory frameworks that ensure medical products remain safe, effective, and high quality.

[334] https://www.fda.gov/science-research/fda-stem-outreach-education-and-engagement
[335] https://commonfund.nih.gov/sites/default/files/OT2-Data-Generation-Projects-B2AI-051321-508.pdf
[336] https://www.fda.gov/about-fda/jobs-and-training-fda/scientific-internships-fellowships-trainees-and-non-us-citizens
[337] https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-2023-public-inventory.csv
[338] https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-2023-public-inventory.csv

3 Healthcare Delivery

3.1 Introduction and Context

U.S. healthcare delivery—defined here as financing, direct provision of patient care, related administrative services and research—is a large and highly complex system. National health expenditures in the U.S. (including public health) were approximately $4.5T in 2022, representing 17% of the U.S. economy and contributing to the employment of approximately 9% of the nation's workforce.[339, 340] In the U.S., healthcare is delivered by licensed providers and predominately financed by payers (e.g., in 2022, 92% of patients in the U.S. had health insurance).[341] A range of stakeholders—beyond patients, providers, and payers—participate in the healthcare delivery ecosystem, including entities that provide resources and technologies that enable care. Many HHS entities, including CMS, HRSA, SAMHSA, IHS, AHRQ, and others, are directly involved in facilitating healthcare delivery or providing guidelines, payment and funding, training, and other operational support to delivery partners.

In healthcare delivery in particular, AI has the potential to enhance a wide range of activities, from care delivery to healthcare finance to research (e.g., health services and behavioral health).[342, 343] HHS aspires to maximize the potential benefit of AI to stakeholders across the healthcare delivery system—to do so, it is essential that AI interventions be patient-centric, with transparency, safety, equity, and security at the forefront of implementation considerations.[344] It is also imperative to protect the safety and security of Americans by ensuring new technology is tested, deployed, and monitored responsibly. In the following chapter, HHS outlines its four goals and actions specific to healthcare delivery: (1) to catalyze health AI innovation and adoption, (2) promote trustworthy AI development and ethical and responsible use, (3) democratize AI technologies and resources, and (4) cultivate AI-empowered workforces and organization cultures.

3.1.1 Action Plan Summary

Later in this chapter, HHS articulates proposed actions to advance its four goals for the responsible use of AI in the sector. Below is a summary of the themes of actions within each goal. For full details of proposed actions please see section 3.6 Action Plan.

[339] https://www.cms.gov/newsroom/fact-sheets/national-health-expenditures-2022-highlights#
[340] https://www.bls.gov/spotlight/2023/healthcare-occupations-in-2022/#
[341] https://www.cms.gov/newsroom/fact-sheets/national-health-expenditures-2022-highlights#
[342] Health services research refers to activities in applied research settings that improve care delivery processes.
[343] https://www.ahrq.gov/healthsystemsresearch/index.html
[344] https://pmc.ncbi.nlm.nih.gov/articles/PMC8826344/#

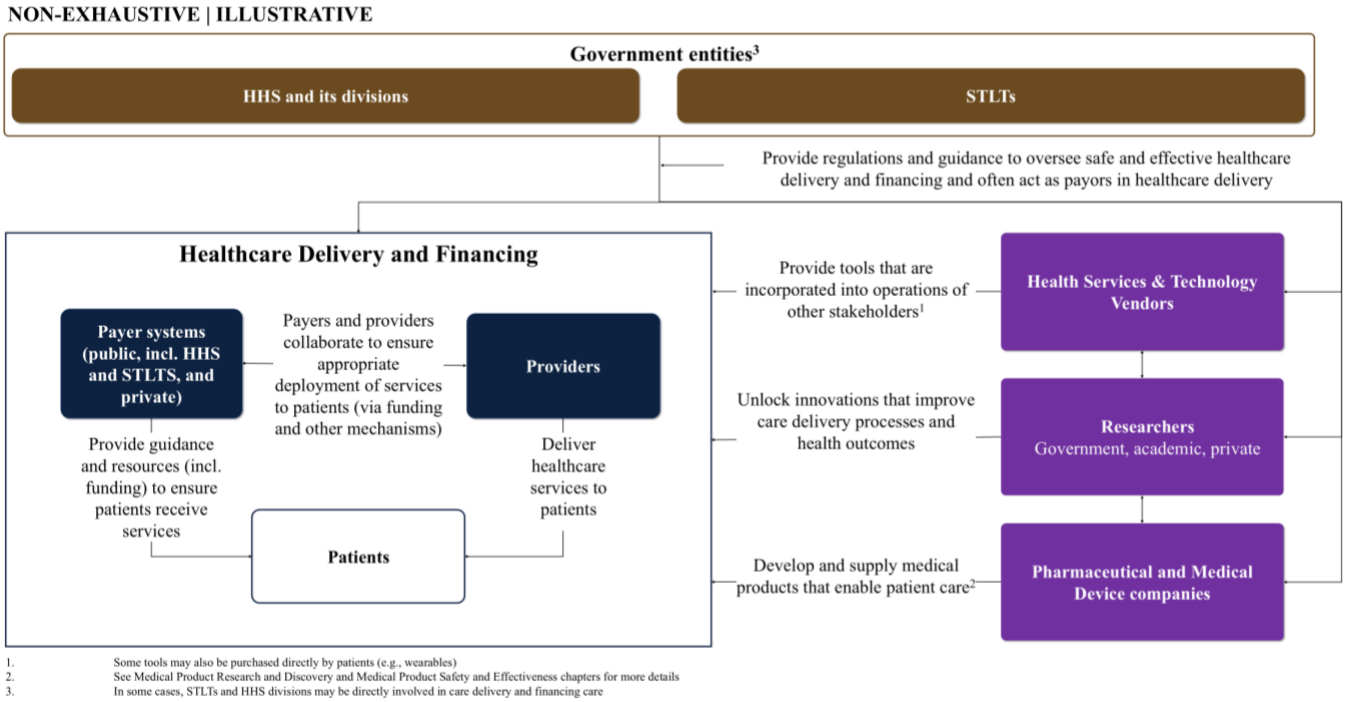
Key goals that actions support	Themes of proposed actions *(not exhaustive, see 3.6 Action Plan for more details)*
1. Catalyzing health AI innovation and adoption	• Supporting the ability to gather evidence for effectiveness, safety, and risk mitigation of AI interventions and best practices for implementation in healthcare delivery settings • Providing guidelines and resources on oversight, medical liability, and privacy and security protections to increase confidence for organizations to develop AI • Ensuring developers and potential deployers of AI have clarity on coverage and payment determination processes to encourage development of AI
2. Promoting trustworthy AI development and ethical and responsible use	• Enhancing enforcement and clarifying guidelines relating to existing requirements • Providing guidelines and support related to organizational governance • Promoting external evaluation, monitoring, and transparency reporting • Enhancing infrastructure to ensure patient safety
3. Democratizing AI technologies and resources	• Promoting equitable access through technical support for and collaboration with delivery organizations that provide services to underserved populations • Providing support for healthcare delivery organizations to address core infrastructure and deployment challenges (i.e., technology, infrastructure, and data infrastructure) that improve AI readiness
4. Cultivating AI-empowered workforces and organization cultures	• Equipping healthcare delivery professionals with access to training, resources, and research to support AI literacy and expertise in their respective health system organizations.

3.2 Stakeholders Engaged in the Healthcare Delivery AI Value Chain

Healthcare delivery is a highly complex set of activities covering the financing of healthcare through public or private health insurance and the provision of healthcare services through private and public hospitals and ambulatory facilities. Employers and individuals purchase healthcare insurance through various entities. Healthcare is delivered by thousands of hospitals and millions of clinicians and other healthcare professionals who offer various services and are regulated by authorities from federal and STLT government entities.

Exhibit 7 shows a non-exhaustive, illustrative diagram of example flows between stakeholders and a bulleted list of stakeholders involved healthcare delivery. Please note that neither the diagram nor the list captures all stakeholder roles and interactions. Please refer to other HHS documents for additional details on regulatory guidance and authorities. Roles may vary depending on healthcare delivery system or activity.

Exhibit 7: Healthcare Delivery Stakeholder Engagement Map



- **HHS divisions and example roles in healthcare delivery (non-exhaustive):**
  - **ACF:** Administers more than 60 programs that provide benefits and services to support families and children, including promoting economic and social well-being. ACF's role in the HHS AI Strategic Plan will focus on ensuring effective and equitable delivery of services to children and families that will promote optimal health.
  - **AHRQ:** Focuses on improving the quality, safety, efficiency, and effectiveness of healthcare for all Americans through research, technology assessments, and work on dissemination and implementation. AHRQ's role in the HHS AI Strategic Plan will focus on promoting and conducting research on the safe adoption of AI that enables high-quality care, disseminating actionable, evidence-based AI knowledge, and provisioning evidence required for coverage decisions.
  - **ARPA-H:** Conducts transformative, high-impact healthcare research across focus areas, including advancing technical solutions, forging a resilient health ecosystem, and driving scalable solutions. ARPA-H's role in the HHS AI Strategic Plan will focus on issuing awards to catalyze cutting-edge research that will improve healthcare delivery.
  - **CDC:** Provides guidelines and research on healthcare delivery for major diseases, supports public health program funding, and may leverage AI to inform and support delivery. CDC's role in the HHS AI Strategic Plan will focus on researching the efficacy of AI in disease prevention and implementing AI in public health efforts.
  - **CMS:** Administers major public healthcare payer programs (e.g., Medicare and Medicaid) and can be involved in setting payment and coverage policies for specific items or services. CMS's role in the HHS AI Strategic Plan will focus on determining coverage and payment of AI-enabled healthcare services, overseeing and certifying state IT systems and data collection standards, and providing technical assistance to providers, states, and other stakeholders. As appropriate, CMS will look to use payment and regulatory policy to ensure trustworthy, responsible use of AI by payers and providers.
  - **FDA:** Helps ensure that human and animal drugs, biological products, and medical devices are safe and effective for their intended uses and that electronic products that emit radiation are safe. As AI becomes a more prominent aspect of medical products and their development, manufacturing operations, and use, the FDA will play a continued role in regulating and supporting stakeholders.

- **HRSA:** Provides equitable healthcare to the nation's highest-need communities, including through programs that support people with low incomes, people with HIV, pregnant women, children, parents, rural communities, transplant patients, and the health workforce. HRSA's role in the HHS AI Strategic Plan will focus on ensuring the equitable use of AI to benefit underserved communities and educating and training future generations of healthcare professionals.
- **IHS:** Provides healthcare services to American Indian and Alaska Native communities. IHS's role in the HHS AI Strategic Plan will focus on implementing healthcare delivery within these populations and ensuring the applicability of AI guidelines to relevant STLTs.
- **NIH:** Supports and conducts biomedical and behavioral research across the U.S. and abroad and can help educate the workforce on AI and promote innovation through its initiatives. NIH's role in the HHS AI Strategic Plan will focus on supporting research on the impact of AI on biomedical and behavioral health, establishing standards in these areas based on research, and unlocking funding to promote the responsible use of AI across HHS service domains.
- **SAMHSA:** Leads public health efforts to advance the behavioral health of the nation and improve the lives of individuals living with mental and substance use disorders, as well as their families. SAMHSA's role in the HHS AI Strategic Plan will focus on providing grant funding and guidelines to STLT communities and collecting, analyzing, and distributing behavioral health data to evaluate programs, improve policies, and raise awareness of resources on prevention, harm reduction, treatment, and recovery.

- **Other federal agencies:** HHS also works closely with many other federal departments, such as the Department of Veterans Affairs and the Department of Housing and Urban Development.
- **Patients, beneficiaries, and their caregivers:** The primary care recipients will interact with the healthcare system as patients in some capacity; in 2020, 83.4% of adults and 94.0% of children reported that they visited a physician or other healthcare provider in the previous year.[345] Caregivers, sometimes serving as guardians, also play a critical role in providing care for infants, children, adolescents, and elder family members.
- **Providers:** These are the primary vehicle for care delivery in the U.S., including:
  - **Healthcare facilities and systems:** The U.S. health system includes approximately 6,100 hospitals (from small community organizations to national systems) in addition to a range of post-acute care settings, outpatient clinics, and long-term care settings.[346, 347]
  - **Clinicians and support staff:** In the U.S. in 2022, there were around 15 million clinical employees, including 933,000 active physicians, 3.4 million registered nurses, and 1.4 million personal care aids, in addition to other clinical staff (e.g., specialists, assistants, therapists, and technicians.[348]
  - **Non-clinical staff:** Non-clinical staff play key roles in organizing and delivering healthcare (e.g., supply chain, maintenance, reception, HR and finance, communications, and IT) and also could engage with AI-enabled tools in administrative settings
  - **Healthcare administration executives:** Medical and health services managers help coordinate and oversee the complex operations of healthcare delivery organizations; 567,200 healthcare administration managers were employed in the U.S. in 2023.[349] Additionally, senior executives, trustees, and boards of directors drive the overarching strategy of delivery organizations and make decisions on AI investments.
- **Payers:** These are public and private organizations that finance patient care and help connect patients to appropriate providers and services based on their needs including:
  - **Public payers (e.g., state Medicaid and other governmental agencies):** Agencies that support implementing regulation, financing, and delivery.
  - **Private payers:** National, regional, and local payers that support financing and care.

[345] https://www.ncbi.nlm.nih.gov/books/NBK587178/

[346] https://data.cms.gov/provider-data/dataset/xubh-q36u More than 5,300 hospitals are registered with Medicare with other care settings making up the balance.

[347] https://www.aha.org/statistics/fast-facts-us-hospitals

[348] https://www.bls.gov/spotlight/2023/healthcare-occupations-in-2022/

[349] https://www.bls.gov/ooh/management/medical-and-health-services-managers.htm

- **Employers:** Employer-sponsored healthcare, which accounts for 54% of managed care lives in the U.S. (often administered by private payers). Employers have an active interest in ensuring the quality and safety of care provided to their employees.[350]

- **STLT governments:** These entities directly perform a variety of healthcare delivery activities, including providing care and financing and providing regulatory oversight of private and public sector activities.
- **Other entities supporting healthcare delivery:**
  - **Technology companies:** A variety of technology vendors actively develop technology for healthcare settings, ranging from diversified, big-tech companies to dedicated healthcare services and technology vendors such as EHRs, revenue cycle management (RCM), and other ancillary services vendors.
  - **Research institutions:** In partnership with healthcare facilities, academic research institutions fuel discoveries that unlock new treatment modalities with the potential to transform the standard of care (e.g., enhanced patient services, new clinical innovations, mitigation of quality and safety issues, newly designed organizational workflows).
  - **Biopharmaceutical and medical device companies:** While the specifics of research and discovery on medical products including drugs, biological products, and devices are covered in other chapters of this plan, these organizations also engage in healthcare delivery via post-launch monitoring, maintenance, and surveillance of AI deployed in clinical settings.
  - **Non-profit and CBOs:** Many of these entities support the direct delivery of referral and care coordination.

3.3 Opportunities for the Application of AI in Healthcare Delivery

AI has the potential to transform care delivery processes, but it also carries inherent risks that must be monitored to ensure positive patient impact and safety. Five ways that AI can support the healthcare system include:

1. **Improving the quality and safety of patient care:** Medical errors, including incorrect and/or delayed diagnoses, may contribute to adverse patient outcomes.[351, 352] AI has the potential to accelerate diagnoses and prevent adverse events by rapidly processing expansive and disparate information, detecting patterns not always apparent to human observation, and directing clinicians to higher-likelihood diagnoses. AI can also enhance care models and health services research to develop innovations that better enable clinicians, payers, and patients.
2. **Improving the patient experience:** AI has the potential to enhance patient satisfaction through more efficient and tailored services that better meet their needs. AI can also provide patients with tools to better understand medical information, including their own medical records and health status, and facilitate more engaged communication with both providers and payers (e.g., through sharing interpretable and relevant patient-facing information).[353]
3. **Automating administrative processes and reduce workforce burden and burnout:** The growth in administrative complexity of healthcare delivery, coupled with shortages in the healthcare workforce, especially in primary care, exacerbates burnout in these already highly demanding work environments.[354, 355, 356] AI applications in administrative contexts – including documentation, member/patient communications, and claims processing - can alleviate resources and provide organizations with more bandwidth to enhance care delivery.

---

[350] https://www.census.gov/content/dam/Census/library/publications/2023/demo/p60-281.pdf

[351] https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/2813854

[352] https://patientsafetyj.com/article/116529-patient-safety-trends-in-2023-an-analysis-of-287-997-serious-events-and-incidents-from-the-nation-s-largest-event-reporting-database

[353] https://pmc.ncbi.nlm.nih.gov/articles/PMC10734361/#section7-20552076231220833

[354] https://www.cms.gov/Outreach-and-Education/Outreach/Partnerships/Downloads/April2019PoPNewsletter.pdf; https://www.healthit.gov/sites/default/files/page/2020-02/BurdenReport_0.pdf

[355] https://bhw.hrsa.gov/data-research/projecting-health-workforce-supply-demand

[356] https://www.ahrq.gov/prevention/clinician/ahrq-works/burnout/index.html#

4. **Enhancing equity and access:** Healthcare disparities are persistent within healthcare, and outcomes can vary by socioeconomic status, location, demographic factors, and more.[357] There is a rapidly growing awareness of the importance of social drivers of health and health-related social needs on health outcomes.[358] AI systems have the ability to incorporate SDOH and other information to inform the identification of at-risk patients, communicate in a patient's preferred language and literacy level, surmount barriers to access for individuals with disabilities, and recommend services and resources better suited to individual circumstances.[359, 360]
5. **Bending the cost curve:** The U.S. remains the highest-cost healthcare system globally, which limits access to care for Americans and hinders U.S. economic productivity. In the aggregate, the adoption of AI across the healthcare delivery value chain could reduce administrative overhead, increase asset and resource utilization, and lessen adverse events,[361] which some reports estimate could reduce annual national healthcare expenditure by up to 10%.[362]

3.4 Trends in AI in Healthcare Delivery

Current trends indicate that the innovative use of AI in healthcare delivery is rapidly evolving. However, there are still barriers to its use. Key trends include:

1. **Investment in health AI is large and growing:** AI accounts for 25% of all healthcare venture capital funding, totaling over $19B since 2021. According to initial reports, roughly two-thirds of this investment has gone into clinical applications of AI and the other third to administrative use.[363]
2. **Mixed enthusiasm and concerns regarding the adoption of AI in the healthcare delivery context:** A recent survey of 100 healthcare executives indicated that over 70% were already pursuing or implementing the technology.[364] However, in another survey, about 40% of physicians indicated they were equally enthusiastic and concerned about using AI.[365] There are concerns that AI adoption could result in a shift in the landscape of healthcare jobs and impact the patient-provider relationship.[366, 367, 368, 369] Patients have similar concerns regarding AI, and results from one survey showed that approximately 60% of respondents were uncomfortable with the possibility of healthcare providers relying on AI.[370, 371] *Additional discussion of these risks and associated actions to mitigation can be found in this chapter's "Action Plan" section.*
3. **Variation in the adoption of AI by healthcare disciplines:** In the 1990s, early uses of ML were applied to medical data to develop the first ML-based systems for diagnosis.[372] AI innovations continue with today's clinical decision support to enable it to be a critical tool for modern clinical workflows. Today, certain applications of AI and ML—particularly in radiology (e.g., reviewing types of medical images such as ECGs, MRI scans, and skin images)—have become widely accepted.[373] While AI applications in radiology have matured, the adoption of AI in other disciplines, like pathology, cardiology, and primary care is

357 https://pubmed.ncbi.nlm.nih.gov/38100101/
358 https://www.cms.gov/priorities/innovation/key-concepts/social-drivers-health-and-health-related-social-needs
359 https://pmc.ncbi.nlm.nih.gov/articles/PMC9976641/
360 https://www.acf.hhs.gov/ai-data-research/artificial-intelligence-acf
361 It is not assumed that AI will eliminate all adverse events.
362 https://www.nber.org/system/files/working_papers/w30857/w30857.pdf
363 https://www.svb.com/trends-insights/reports/artificial-intelligence-ai-in-healthcare/
364 https://www.mckinsey.com/industries/healthcare/our-insights/generative-ai-in-healthcare-adoption-trends-and-whats-next#/ Survey where executives from 100 healthcare organizations were surveyed on their intentions to implement GenAI.
365 https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf
366 https://hbr.org/2019/10/ai-can-outperform-doctors-so-why-dont-patients-trust-it
367 https://www.fastcompany.com/91053431/surveys-show-americans-dont-trust-ai-medical-advice-why-that-matters
368 https://insight.kellogg.northwestern.edu/article/will-ai-replace-doctors
369 https://pmc.ncbi.nlm.nih.gov/articles/PMC10811613/
370 https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/
371 https://hbr.org/2019/10/ai-can-outperform-doctors-so-why-dont-patients-trust-it
372 https://www.nejm.org/doi/full/10.1056/NEJM199406233302512
373 https://www.nejm.org/doi/full/10.1056/NEJMra2302038

growing.[374, 375] *Additional analyses of use cases can be found in this chapter's "Use Cases and Risks" section.*

4. **Increased innovation and uptake of administrative AI use:** AI use in administrative tasks has advanced over the last few years, given lower development costs compared to clinical use cases and the onset of GenAI and LLM technology.[376] Recent applications include "extract[ing] drug names from physicians' notes, reply[ing] to patient administrative questions, summariz[ing] medical dialogues, and writ[ing] histories and physical assessments."[377] According to an American Medical Association survey, 54% of physicians are enthusiastic about using AI in their practices (particularly for administrative tasks such as documentation and charting).[378]
5. **Heterogeneity in organizations' data and technology systems:** The variation that exists in healthcare organizations' access to technology and resources needed to use AI—including data management, clinical and administrative applications, and core infrastructure (e.g., cloud computing)—impacts current adoption.[379] Heterogeneity in data modalities (e.g., numerical, textual, images, video, and audio) and standards across healthcare systems and EHRs create additional barriers to AI applications across the healthcare sector.[380] This heterogeneity also contributes to organizations' decision-making on which solutions to build, partner with (e.g., with AI vendors),[381] or procure from others, and to what degree (e.g., AI, GenAI, or non-AI interventions).[382, 383, 384]

3.5 Potential Use Cases and Risks for AI in Healthcare Delivery

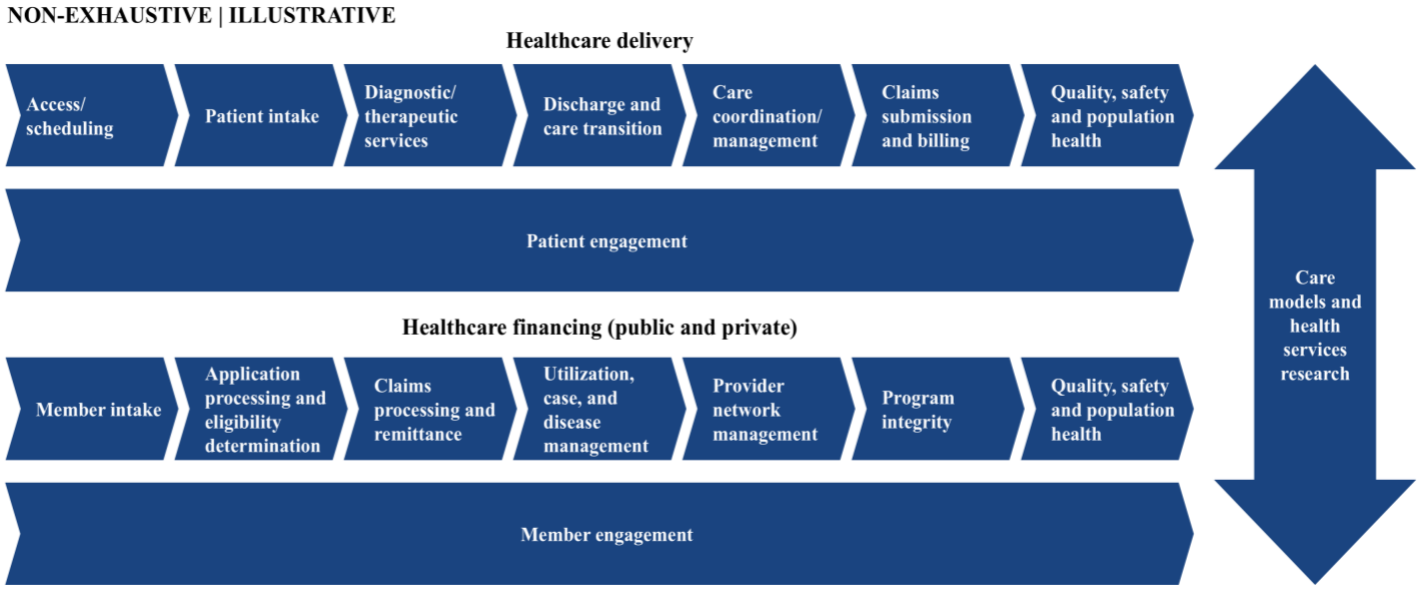
Healthcare delivery and financing include a wide range of activities, all of which are likely to be impacted by existing and emerging AI, though some may be more impacted than others. The use of AI in healthcare delivery and financing—as contemplated in this chapter—can be considered across the value chain of activities in healthcare delivery (e.g., diagnostic services, patient care delivery), financing (e.g., claims processing, provider network management), and research (see Exhibit 8).

There is variation in the type of technology and complexity across AI use cases (e.g., simpler rule-based automation versus complex LLMs), and thus, some have higher rates of adoption across the health system relative to others that are in more nascent stages of testing.

There is also a broad range of risks posed by AI within healthcare delivery, including an impact on patient safety, deterioration of patient-provider relationships, and barriers to or inappropriate administration of care resulting from algorithmic bias. As discussed earlier in the document, HHS and its divisions (e.g., CMS) provide frameworks to both consider and mitigate risks in healthcare AI, such as FAVES.

374 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10487271/
375 https://pmc.ncbi.nlm.nih.gov/articles/PMC10517477/
376 https://www.svb.com/globalassets/trendsandinsights/reports/svb-the-ai-powered-healthcare-experience-2024.pdf
377 https://jamanetwork.com/journals/jama/fullarticle/2808296
378 https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf
379 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8285156/
380 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9908503/#
381 https://healthinnovation.ucsd.edu/news/11-health-systems-leading-in-ai
382 https://pmc.ncbi.nlm.nih.gov/articles/PMC9628307/#
383 https://scopeblog.stanford.edu/2019/02/26/ai-will-not-solve-health-care-challenges-yet-but-there-are-innovative-alternatives-researcher-writes/
384 https://www.nejm.org/doi/full/10.1056/NEJMra2204673

Exhibit 8: Healthcare Delivery and Financing Value Chains



3.5.1 AI in Delivery

While the individual activities provided by a provider organization will vary greatly in size and focus (e.g., primary care clinics, large multispecialty groups, academic medical centers, state agencies, and federally qualified health centers), the value chain is intended to describe the core set of healthcare delivery functions that frequently apply and the potential benefits or applications of AI.

Innovation, development, and uptake of AI are inconsistent across the value chain—they are relatively more advanced in administrative functions (including those with clinical and non-clinical impact, such as operating room optimization, call-center enablement, talent management, and back-office administration), while AI applications in diagnostics and therapeutic services are still less common outside of radiology.

Overall, AI has had relatively higher levels of adoption in use cases where data is readily available (e.g., through EHRs or wearable devices), and is still nascent in applications for complex cases with limited data availability (i.e., given risks of model inaccuracy or bias toward specific populations).[385, 386] Areas such as care coordination and transitions that require connecting disparate data sources (e.g., remote monitoring and hospital and home-care records) could be ripe for opportunity, but they continue to be limited in adoption, given challenges in connecting underlying data.

Larger hospitals are further along in AI uptake, whereas smaller hospitals and physician groups are near the beginning of their AI journeys, piloting some AI use cases. However, as discussed previously, the relative value of certain AI use cases will vary based on an individual organization's characteristics (e.g., provider size, needs, resources, existing capabilities, and service areas).

In the tables below, HHS highlights a non-exhaustive list of potential benefits and risks of AI across the healthcare delivery value chain. Please note that the use cases detailed below highlight existing or potential ways that AI can be used by a variety of stakeholders in this domain. For details on how HHS and its divisions are using AI, please

[385] https://pmc.ncbi.nlm.nih.gov/articles/PMC7979747/#
[386] https://pmc.ncbi.nlm.nih.gov/articles/PMC7414411/#

reference the HHS AI Use Case Inventory 2024.[387] Further, use-cases and risks related to financing and research are discussed in 3.5.2 AI in Financing and 3.5.3 AI in Care Models and Health Services Research, respectively.

Functional component 1: Access and/or scheduling

The process of scheduling patients for appointments and services

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Streamlined and automated scheduling tools to optimize efficiency *E.g., predictive analytics to reduce no-shows* Targeted interventions (e.g., outreach) can substantially increase show rates for patients most likely to miss appointments.[388, 389] *E.g., appointment scheduling optimization* AI can optimize scheduling by predicting patient appointment preferences and availability, reducing wait times, and improving clinic efficiency.[390] *E.g., operating room scheduling optimization* AI can analyze surgical schedules, patient data, and resource availability to optimize operating room usage, reducing downtime and improving surgical throughput.[391]	**Potential to introduce bias** *E.g., mismatched overbooking of appointments* Applying a one-size-fits-all approach to overbooking appointments based on no-show rates may disproportionately impact patients with certain characteristics (e.g., socioeconomic status, low access to transportation, and fear of doctors or hospitals).[392, 393] *E.g., over-emphasis of variables that enhance disparities in scheduling* AI use for procedure scheduling (e.g., operating room scheduling) could risk perpetuating disparities in access to care if algorithms trained on current resource allocation data give too much weight to certain variables (e.g., procedure profitability, coverage type).[394]

Functional component 2: Patient intake and support

The initial stage of gathering and verifying patient information, including medical history and insurance details, to prepare for treatment and ensure smooth administrative processes

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Personalized AI-assisted patient intake processes to increase efficiency and patient satisfaction *E.g., streamlined patient data collection* Auto-generation and tracking of communications sent to patients to minimize duplicate data collection and patient burden[395] *E.g., automated AI voice technology* AI-driven conversational voice technology to automate patient intake processes (e.g., through recording and transcription)[396]	**Potential to magnify patient trust concerns** *E.g., overcollection of patient data* The overcollection of data (or perception of data misuse, even if inaccurate) for AI models can cause patient discomfort in care delivery processes and create or enhance distrust, particularly for populations who may already have negative perceptions of the healthcare system.[397]

[387] https://www.healthit.gov/hhs-ai-usecases
[388] https://www.healthcareitnews.com/news/fqhc-slashed-its-patient-no-show-rate-ai-3-months
[389] https://pmc.ncbi.nlm.nih.gov/articles/PMC10150669/
[390] https://pmc.ncbi.nlm.nih.gov/articles/PMC10905346/#
[391] https://www.nejm.org/doi/full/10.1056/NEJMra2204673
[392] https://pmc.ncbi.nlm.nih.gov/articles/PMC7280239/pdf/rmhp-13-509.pdf
[393] https://www.healthaffairs.org/content/forefront/discrimination-artificial-intelligence-commercial-electronic-health-record-case-study
[394] https://www.healthaffairs.org/content/forefront/discrimination-artificial-intelligence-commercial-electronic-health-record-case-study
[395] https://www.nber.org/system/files/working_papers/w30857/w30857.pdf
[396] https://pubmed.ncbi.nlm.nih.gov/33999834/
[397] https://www.ama-assn.org/system/files/ama-patient-data-privacy-survey-results.pdf

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Automated tools to reduce administrative tasks and free up staff to focus on patient care and more complex issues *E.g., optimized patient request handling* AI virtual agents can quickly answer simple patient requests (e.g., as one health system did during the COVID-19 pandemic by using an NLP-driven chatbot to direct a large influx of patient calls to the appropriate system to facilitate their requests).[398]	**Potential to impede patient access to care** *E.g., incorrect decisions enabled by AI based on patient data* Erroneous data collected by AI could lead to inappropriate decisions and denial of services.

Functional component 3: Diagnostic/therapeutic services

The delivery of medical care, including diagnosis and treatment, is supported by advanced systems like EHR, clinical decision support, wearables, and telehealth tools to improve the quality and efficiency of care

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Automated documentation and summarization of patient information to increase healthcare worker efficiency[399] *E.g., ambient listening* AI-driven ambient listening systems can capture and transcribe patient-provider interactions in real time, facilitating more accurate documentation and diagnosis and enabling providers to focus more on patient care and improving the patient experience.[400, 401]	**Potential for inappropriate application** *E.g., confabulations* Automated documentation systems may generate false information on a patient's medical history and lead to inappropriate care recommendations, underscoring the importance of human-in-the-loop and robust confabulation detection methods.[402] *E.g., AI impacting patient-clinician relationships and trust* Patients have expressed concerns that utilizing AI for clinical decision-making may deteriorate patient-provider relationships, as AI continually automates tasks typically done by humans—especially given the emotional and personal nature of experiencing medical conditions, underscoring the importance of empathetic and compassionate interactions within healthcare delivery.[403, 404]

[398] https://www.nejm.org/doi/full/10.1056/NEJMra2204673
[399] https://catalyst.nejm.org/doi/full/10.1056/CAT.23.0404
[400] https://med.stanford.edu/news/all-news/2024/03/ambient-listening-notes.html
[401] https://www.ama-assn.org/system/files/2019-01/augmented-intelligence-policy-report.pdf
[402] https://openreview.net/pdf?id=6eMIzKFOpJ
[403] https://pmc.ncbi.nlm.nih.gov/articles/PMC10116477/#
[404] https://journalofethics.ama-assn.org/article/how-will-artificial-intelligence-affect-patient-clinician-relationships/2020-05

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Automated intelligence tools to support the evaluation of diagnosis and treatment options and surface critical insights about patient conditions *E.g., prediction and risk identification* AI algorithms can analyze patient health indicators to predict disease outcomes (e.g., one health system used an AI algorithm to predict sepsis in patients by combining EHR data with blood pressure and heart rate measures).[405, 406] *E.g., precision medicine* AI can power CDS tools to help physicians consider optimal interventions and help surface critical (and potentially challenging to trace) insights about patient conditions.[407]	**Potential misuse or misinterpretation of health data** *E.g., ineffective treatment plans informed by AI* Potential prioritization of testing data and analysis over patient-reported indicators and other factors in AI-generated behavioral health treatment decision support could lead to misdiagnoses and treatments that may worsen behavioral health outcomes and trust.[408] *E.g., health technologies may not consider nuances of individuals* AI tools may not account for demographic and SDOH factors such as communication barriers, which may increase technological concerns among patients and lead to reduced patient satisfaction, trust, and effectiveness in care.[409]
Analysis of patient data to develop targeted interventions or educational materials *E.g., sentiment analysis through multiple data formats* AI can process unstructured data (e.g., text posted on social media, user input) to generate summaries of perspectives on mental health. These can be used to develop and disseminate personalized educational materials, guidance, strategies, and referrals.[410] *E.g., AI analysis to develop combined interventions* AI can help create holistic treatment plans that combine multiple types of interventions (e.g., behavioral and clinical interventions) such as guided diet monitoring and AI-tailored education paired with CDS (e.g., measurement of health indicators and personalized medication plans) for diabetes patients.[411] *E.g., AI technology that provides reminders and measures medicine intake* AI tools such as smartphone apps can assess and encourage adherence through daily monitoring and reminders (e.g., smartphone camera to confirm ingestion of drug).[412]	

[405] https://ai.nejm.org/doi/full/10.1056/AIp2300031 Use of GPT-4 to Diagnose Complex Clinical Cases

[406] https://www.sciencedirect.com/science/article/abs/pii/S1553725020300969?via%3Dihub

[407] https://www.mcpdigitalhealth.org/article/S2949-7612(24)00041-5/fulltext

[408] https://www.aha.org/aha-center-health-innovation-market-scan/2024-05-14-will-ai-help-address-our-behavioral-health-crisis

[409] https://pmc.ncbi.nlm.nih.gov/articles/PMC8521858/#

[410] https://pmc.ncbi.nlm.nih.gov/articles/PMC10982476/#

[411] https://pmc.ncbi.nlm.nih.gov/articles/PMC10591058

[412] https://pmc.ncbi.nlm.nih.gov/articles/PMC8521858/#s3

Functional component 4: Discharge and care transition

Managing the process of transitioning patients from one care setting to another, ensuring continuity of care and proper follow-ups through integrated systems and patient engagement platforms

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
AI algorithms that analyze patient circumstances and enable more personalized and efficient care transition processes *E.g., patient-facing virtual care assistants* AI can increase education and transparency by explaining a diagnosis and care management plan, giving patients a 24/7 resource that educates them and provides timely information through a virtual care assistant or chatbot.[413, 414] *E.g., chatbots that minimize potential engagement with clinicians* AI chatbots can help encourage and deliver care for patients who may have conditions they perceive as embarrassing or stigmatizing and would prefer not to have an in-person consultation.[415]	**Potential for inappropriate application** *E.g., confabulation of inappropriate recommendations* AI models can make errors in data analysis, incorrectly transcribe recordings, or convey false information to clinicians.[416] *E.g., deterioration of key skillsets* Additional introduction of AI tools may result in over-reliance on these technologies by clinicians, potentially leading to deskilling in nuanced areas of health, particularly where human empathy and engagement play a significant role.[417]

Functional component 5: Care coordination and management

Ongoing management of patient care across different services and providers, utilizing digital tools and analytics to enhance care coordination, patient engagement, and overall health outcomes

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Remote monitoring of patient conditions to enhance patient care effectiveness and timeliness *E.g., chronic care management* AI decision aids can support ongoing disease management by providing patients with tools that support reminders, predict issues, and flag care needs to providers and patients. Additionally, they can be used in hospital settings to monitor care across disease areas (e.g., glucose changes for someone with diabetes but who is hospitalized for other acute needs).[418, 419, 420, 421]	**Potential to introduce bias** *E.g., incorrect risk stratification by demographic* AI algorithms used in care coordination decision-making may be vulnerable to bias by assigning the same level of risk to patients despite characteristics that should be taken into consideration to determine risk (e.g., one AI algorithm used by a health system reduced the number of minority patients identified for care, even though that cohort of patients was sicker and needed more care).[422, 423]

413 https://pubmed.ncbi.nlm.nih.gov/37054749/, https://pmc.ncbi.nlm.nih.gov/articles/PMC10219811/

414 https://cdsic.ahrq.gov/sites/default/files/2024-09/PAIGE%20Assessment%20Report_Public%20Version.pdf

415 https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2023.1275127/full

416 https://openreview.net/pdf?id=6eMIzKFOpJ

417 https://www.sciencedirect.com/science/article/pii/S2949916X24000938#

418 https://www.jmir.org/2023/1/e42335/PDF

419 https:/pubmed.ncbi.nlm.nih.gov/38215713 Remote Monitoring and Artificial Intelligence: Outlook for 2050.

420 https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/widm.1485

421 https://www.annallergy.org/article/S1081-1206(21)01276-X/abstract Methods to engage patients in the modern clinic.

422 https://www.nature.com/articles/s41746-023-00858-z# Bias in AI models for medical applications: challenges and mitigation strategies.

423 https://www.science.org/doi/10.1126/science.aax2342

Functional component 6: Claims submission and billing

Submitting claims for reimbursement and managing billing are often automated to ensure timely and accurate payment, reduce denials, and optimize RCM

Potential benefits and example use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Tools to help measure and assist physicians in choosing and logging optimal interventions[424, 425] *E.g., billing code automation and analysis* Automating billing codes and checking the accuracy of billing based on unstructured notes and data[426, 427]	**Potential for increased barriers to patient care** *E.g., inaccurate claims submissions* Inaccurate claims submissions caused by AI may occur due to model failures (e.g., poor/exposed data, analysis methodology, interpretation) and lead to increased liability for medical professionals and fines.[428, 429] *E.g., expanding costs due to competition in payment integrity/ revenue cycle management* As providers invest in AI to optimize revenue and payers invest in AI to increase payment integrity, the potential for meaningful costs to the system increases—with the additional risk of affecting patients.[430, 431, 432]

[424] https://www.medicaleconomics.com/view/revolutionizing-denials-management-with-artificial-intelligence

[425] https://www-nejm-org.ezproxyhhs.nihlibrary.nih.gov/doi/10.1056/NEJMra2204673

[426] https://www-nejm-org.ezproxyhhs.nihlibrary.nih.gov/doi/10.1056/NEJMra2204673

[427] https://www.medicaleconomics.com/view/revolutionizing-denials-management-with-artificial-intelligence; https://www-nejm-org.ezproxyhhs.nihlibrary.nih.gov/doi/10.1056/NEJMra2204673

[428] https://link.springer.com/article/10.1007/s40273-019-00777-6#

[429] https://oig.hhs.gov/compliance/physician-education/fraud-abuse-laws/#

[430] https://www.hfma.org/revenue-cycle/denials-management/health-systems-start-to-fight-back-against-ai-powered-robots-driving-denial-rates-higher/

[431] https://jamanetwork.com/journals/jama/fullarticle/2812255 AI Alone Will Not Reduce the Administrative Burden of Healthcare

[432] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2816204 Denial—Artificial Intelligence Tools and Health Insurance Coverage Decisions.

Functional component 7: Quality, safety, and population health

Ensuring healthcare services meet established standards of quality and safety, using tools like AI-powered support, decision support systems, and continuous monitoring to improve clinical care and organizational performance

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
AI tools to enhance patient care and hospital quality measures *E.g., adverse event and re-admission prevention* AI can remotely monitor patient conditions to prevent re-admissions by identifying risks of potential deterioration and prioritizing interventions, ensuring timely and effective care.[433] *E.g., quality measurement* Abstraction and analytics tools for more accurate and efficient hospital quality measurement[434]	**Potential for bias** *E.g., underrepresentation of certain populations in training data* Underlying training data may be biased due to historical disparities in access and quality of care delivery.[435]

3.5.2 AI in Financing

The financing landscape features a wide variety of payers (e.g., Medicare, state Medicaid agencies, large national insurers, regional specialty payers, and managed care organizations). There are key variations among these organizations (e.g., populations served) and in their payment structures (e.g., value-based care, fee-for-service). Across these, there are wide ranges of use cases and risks for these payers, which include the examples listed in the table below.[436, 437]

In financing, AI and LLMs are increasingly being used for a range of functions and tasks, including prior authorization, clinical review assessments, utilization management, and claims adjudication.[438] Given the industry's extensive data analytics and document processing, a large and expanding wave of new use cases is expected in the coming years. Expansion in this segment is not without challenges: there have been ongoing litigation and concerns from Congress regarding the use of AI and algorithms to deny prior authorization requests, particularly in how AI complies with state and federal regulations impacting payer decision-making.[439]

[433] https://www.nejm.org/doi/full/10.1056/NEJMra2204673
[434] https://ai.nejm.org/doi/full/10.1056/AIcs2400420
[435] https://pmc.ncbi.nlm.nih.gov/articles/PMC10497548/#CR49
[436] https://www.nejm.org/doi/full/10.1056/NEJMra2204673
[437] https://www.nber.org/system/files/working_papers/w30857/w30857.pdf
[438] https://www.healthaffairs.org/content/forefront/ai-and-health-insurance-prior-authorization-regulators-need-step-up-oversight
[439] https://www.statnews.com/2023/11/14/unitedhealth-class-action-lawsuit-algorithm-medicare-advantage/

Functional component 1: Member intake

The process of enrolling individuals in a healthcare insurance plan, ensuring their information is accurately captured and maintained for future interactions

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Streamlined enrollment tools that personalize member engagement *E.g., generating personalized member outreach* AI can analyze member data to create tailored communication strategies (e.g., through GenAI) that address specific health needs, preferences, and engagement patterns.	**Potential for privacy concerns** *E.g., over-personalized outreach autogenerated by AI* Communications may be perceived as intrusive, and AI over-personalization could be perceived as the overcollection or overuse of member data.[440]

Functional component 2: Application processing and eligibility determination

Reviewing and verifying applications to determine if applicants meet the criteria for coverage, ensuring that only eligible individuals receive benefits

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
AI to significantly reduce manual application-centric workloads *E.g., application review* AI can support the rapid review of applications to identify missing information, check other eligibility (e.g., secondary coverage), and support other functions. *E.g., adaptive customer-facing chatbots* AI-driven chatbots can be trained to handle a large variety of inquiries—from eligibility questions to application status updates—providing instant and accurate responses to members and reducing call center burden.[441]	**Potential to introduce bias** *E.g., incorrect denial of eligibility* Without proper calibration or human-in-the-loop, models risk denying eligibility—particularly to populations with historically more complicated coverage—creating significant hurdles for patients to receive necessary procedures.[442] *E.g., exacerbating underserved populations' distrust of care* Inaccurate responses to populations already less likely to seek care and support may further discourage care-seeking behavior.[443]

[440] https://www.ama-assn.org/system/files/ama-patient-data-privacy-survey-results.pdf
[441] https://www.ncbi.nlm.nih.gov/books/NBK602381/
[442] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2816204
[443] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2816204

Functional component 3: Claims processing and remittance

Handling and adjudicating claims submitted by healthcare providers, ensuring timely payment or denial based on policy terms and services rendered

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Automatic AI processing of complex claims data to streamline decision-making *E.g., fast-tracking claims approvals* Predictive analytics can generate summaries and rapidly assess the validity of claims to fast-track approvals.[444] *E.g., automated claims review* AI algorithms can automate the review of claims for errors, inconsistencies, and compliance with policy terms, speeding up the process and reducing manual effort.[445]	**Potential to exacerbate costs in the system** *E.g., expanding costs due to competition in payment integrity/RCM* As providers invest in AI to optimize their revenues and payers invest in AI tools to increase their payment integrity capabilities, incremental costs could occur through administrative waste—this could affect patients (hospitals billing patients in case of denials by payers).[446, 447, 448]

Functional component 4: Utilization, case, and disease management

Monitoring and managing the use of healthcare services to ensure they are necessary and cost effective, thereby optimizing resource use and controlling costs

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive analytic tools to optimize healthcare service delivery to patients *E.g., patient re-admission analysis and prevention* AI interventions can support case management programs by predicting which patients are at higher risk and supporting targeted interventions to prevent future re-admissions (in one example, AI interventions reduced re-admission rates by 55%).[449] **Automatic AI processing of utilization management and prior authorization** *E.g., prior authorization adjudication* AI can streamline the prior authorization process by quickly verifying necessary medical information and automating approval workflows, reducing delays in patient care.[450,451]	**Potential to generate inappropriate outcomes** *E.g., AI decision support tools used in coverage determinations* AI decision support tools used to support determination of coverage for services may be inconsistent with terms of coverage, a specific patient's circumstances, or fail to abide by applicable federal or state law.[452]

444 https://pmc.ncbi.nlm.nih.gov/articles/PMC6616181/

445 https://www.nber.org/system/files/working_papers/w30857/w30857.pdf

446 https://www.hfma.org/revenue-cycle/denials-management/health-systems-start-to-fight-back-against-ai-powered-robots-driving-denial-rates-higher/

447 https://jamanetwork.com/journals/jama/fullarticle/2812255 AI Alone Will Not Reduce the Administrative Burden of Healthcare.

448 https://jamanetwork.com/journals/jama-health-forum/fullarticle/2816204 Denial—Artificial Intelligence Tools and Health Insurance Coverage Decisions

449 https://www.nejm.org/doi/full/10.1056/NEJMra2204673

450 https://www.nejm.org/doi/full/10.1056/NEJMra2204673

451 https://pubmed.ncbi.nlm.nih.gov/36809561/ Could an artificial intelligence approach to prior authorization be more human?

452 https://www.aha.org/system/files/media/file/2024/02/faqs-related-to-coverage-criteria-and-utilization-management-requirements-in-cms-final-rule-cms-4201-f.pdf

Functional component 5: Provider network management

Managing relationships and contracts with healthcare providers to ensure a robust and effective network for members that facilitates access to necessary services

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Algorithms that compare quantitative network metrics (e.g., rates, credentialing compliance) to identify areas of variability and potential for standardization *E.g., provider rate comparison* AI can compare rates across different providers and services, helping payers and patients make informed decisions about cost-effective care options and negotiate better rates.[453] **Algorithms that streamline documentation processes to support expansive provider network** *E.g., automated provider credentialing* AI can automate provider credential verification, ensuring that all necessary qualifications and certifications are up to date and reducing the administrative burden on healthcare organizations.[454]	**Potential to exacerbate bias** *E.g., increased bias caused by AI algorithms* Safety-net hospitals, which are typically low-margin and care for underrepresented populations, may be further disadvantaged in payer negotiations with payors using sophisticated AI algorithms to manage their network strategy.[455, 456]

Functional component 6: Program integrity

Implementing measures, including advanced analytics, to prevent fraud, waste, and abuse within the healthcare insurance system to ensure the integrity and sustainability of the program

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Algorithms that detect and mitigate fraud to protect patients *E.g., fraud detection* Provider-ranking algorithms can identify fraud using a corpus of publicly and privately available data. CMS has launched a Fraud Prevention System that uses predictive analytics to screen claims before payment, using indicators that flag fraud and enable protective interventions.[457]	**Potential for unintended consequences or inappropriate outcomes** *E.g., increased financial burden* Payer investment in AI tools that increase adverse coverage decisions may financially impact patients and organizations as hospitals increase billing to offset revenue lost from increased denials.[458, 459, 460]

[453] https://www.nber.org/system/files/working_papers/w30857/w30857.pdf

[454] https://www.beckershospitalreview.com/strategy/the-role-of-ai-in-clinician-credentialing-and-enrollment-a-balanced-perspective.html

[455] https://www.chcf.org/wp-content/uploads/2024/04/ExaminingAIandHealthCare.pdf

[456] https://www.science.org/doi/10.1126/science.aax2342 Dissecting racial bias in an algorithm used to manage the health of populations.

[457] https://www.cms.gov/About-CMS/Components/CPI/Widgets/Fraud_Prevention_System_2ndYear.pdf

[458] https://www.hfma.org/revenue-cycle/denials-management/health-systems-start-to-fight-back-against-ai-powered-robots-driving-denial-rates-higher/

[459] https://jamanetwork.com/journals/jama/fullarticle/2812255 AI Alone Will Not Reduce the Administrative Burden of Healthcare.

[460] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2816204 Denial—Artificial Intelligence Tools and Health Insurance Coverage Decisions.

Functional component 7: Quality, safety, and population health

Ensuring that healthcare services provided to members meet established standards of quality and safety, including continuously monitoring and improving these standards

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
AI models that monitor patient and provider indicators (e.g., sentiment analysis) to gather feedback and improve care quality *E.g., patient experience analysis* AI can analyze patient feedback from surveys and other sources to identify satisfaction, trends, and areas for improvement in care, providing actionable insights to payers about quality of care within their provider network.[461]	**Potential to introduce bias** *E.g., failure to identify diseases in patient populations* AI algorithms trained on specific patient populations may be biased, leading to inaccurate conclusions regarding patient safety (e.g., a sepsis prediction algorithm built on a hospital's EHR only identified the condition in 7% of the patient population, delaying care for others in need and inaccurately representing quality of patient care).[462]

[461] https://www.nejm.org/doi/full/10.1056/NEJMra2204673

[462] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2815239

3.5.3 AI in Care Models and Health Services Research

The use of AI in care model and health services development is growing within applied research settings. AI also informs the development of non-device behavioral interventions (e.g., cognitive behavioral therapy, nutrition counseling), which can lead to the generation, modification, adaptation, or refinement of existing interventions.[463, 464] HHS divisions support research and innovation in these areas, including AHRQ (e.g., AI and safety NOFO,[465] guidance on mitigating algorithmic bias),[466] CMS (e.g., CMMI outcomes challenge), NIH (e.g., Office of Behavioral and Social Sciences Research),[467] and SAMHSA (e.g., Center for Behavioral Health Statistics and Quality).[468]

Care model and health services research

Analyzing and optimizing healthcare delivery, workforce models, financial performance, and patient outcomes through innovative, data-driven, and value-based approaches to improve health system performance and equity.

Note: This refers to AI-based research into care models, which may involve medical products but pertains primarily to the use of AI to improve healthcare delivery. Discussion pertaining to medical product development is found in other chapters, notably Medical Product Development, Safety, and Effectiveness.

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Clinical pathway and care model optimization/generation *E.g., AI-generated care pathways* AI can recommend and optimize clinical care pathways, which ensures that patient care aligns with evidence-based guidelines and reduces variation in clinical care between practitioners.[469] *E.g., population-data-enhanced care models* AI can synthesize large volumes of data and generate customized care models. By aligning incentives around patient outcomes, AI can help payers develop value-based care models. Predictive analytics can identify trends and help develop care models for patients.[470, 471] *E.g., digital twins to measure patient conditions* AI can analyze patient data from various sources, including EHRs, wearables, medical devices, and more, to generate digital twins that help provide early detection of health risks and create proactive interventions.[472, 473]	**Potential to introduce bias** *E.g., inaccurate conclusions in research on patient populations* AI algorithms trained on specific patient populations may be biased, leading to misrepresentative findings from research that do not apply equally across groups or perpetuate existing biases.[474, 475]

[463] https://www.nia.nih.gov/research/dbsr/nih-stage-model-behavioral-intervention-development
[464] https://www.samhsa.gov/resource/dbhis/trauma-focused-cognitive-behavioral-therapy-tf-cbt
[465] https://grants.nih.gov/grants/guide/pa-files/PA-24-261.html
[466] https://www.ahrq.gov/news/newsroom/press-releases/guiding-principles.html
[467] https://obssr.od.nih.gov/
[468] https://www.samhsa.gov/about-us/who-we-are/offices-centers/cbhsq
[469] https://healthsciencepub.com/index.php/jaihm/article/view/88/84
[470] https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-023-04698-z
[471] https://pmc.ncbi.nlm.nih.gov/articles/PMC11269274/
[472] https://pubmed.ncbi.nlm.nih.gov/31649194/
[473] https://ai.cms.gov/assets/CMS_AI_Playbook.pdf
[474] https://pmc.ncbi.nlm.nih.gov/articles/PMC6347576/
[475] https://postgraduateeducation.hms.harvard.edu/trends-medicine/confronting-mirror-reflecting-our-biases-through-ai-health-care

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Predictive tools informing research into patient outcomes and care models *E.g., AI-enabled smartphone applications for medication adherence* An AI-enabled smartphone app can provide reminders and dosage instructions and then confirm ingestion to detect non-adherence and predict future non-adherence.[476] This data and any research findings from this work can be used to inform direct care and design of care models. *E.g., predictive rapid response system for in-hospital cardiac arrest* AI-based algorithm for predicting events of deterioration (e.g., cardiac arrest and unexpected ICU admission), which could be used to improve decision-making and design of care models.[477]	

3.6 Action Plan

In light of the evolving AI landscape in healthcare delivery, HHS has already taken multiple steps including issuance of new guidelines and rules and launch of health AI related programs to promote responsible AI. The Action Plan below follows the four goals that support HHS's AI strategy: 1. catalyzing health AI innovation and adoption; 2. promoting trustworthy AI development and ethical and responsible use; 3. democratizing AI technologies and resources; and 4. cultivating AI-empowered workforces and organization cultures. For each goal, the Action Plan provides context, an overview of HHS and relevant other federal actions to date, and specific near- and long-term priorities HHS will take. HHS recognizes that this Action Plan will require revisions over time as technologies evolve and is committed to providing structure and flexibility to ensure longstanding impact.

3.6.1 Catalyze Health AI Innovation and Adoption

HHS has an opportunity to increase AI innovation and adoption safely through the following actions:

1. Supporting the ability to gather evidence for effectiveness, safety, and risk mitigation of AI interventions and best practices for implementation in healthcare delivery settings
2. Providing guidelines and resources on oversight, medical liability, and privacy and security protections to increase confidence for organizations to develop AI
3. Ensuring developers and potential deployers of AI have clarity on coverage and payment determination processes to encourage development of AI

Below, HHS discusses the context, HHS and other federal actions to date, and plans to catalyze health AI innovation and adoption in healthcare delivery.

1. **Supporting the ability to gather evidence for effectiveness, safety, and risk mitigation of AI interventions and best practices for implementation in healthcare delivery settings**

   Context:

   There is variation in both confidence and understanding of AI and concerns about its potential impacts among clinicians and other leaders in delivery settings. Some disciplines, such as radiology, have a more established track record of working with AI in clinical settings. In contrast, others are less likely to see AI applications beyond administrative settings in the present state. According to an AMA survey, 56% of physicians believe

[476] https://pmc.ncbi.nlm.nih.gov/articles/PMC8521858/
[477] https://pubmed.ncbi.nlm.nih.gov/32205618/

the most promising AI use cases are in supporting administrative tasks.[478] Further research on the application of AI in complex clinical settings could unlock innovation and incentivize adoption by providing an evidence-based foundation for the appropriate and safe use of AI. These efforts could also aim to build evidence to address clinicians' and other stakeholders' concerns to ensure that AI is adopted in ways most helpful to patients and those engaged in their care. Such an approach will also help sustain effective and responsible use of AI by building confidence in these technologies for patients, clinicians, and other stakeholders based on an informed understanding of their benefits.

Given that healthcare organizations in the U.S. are highly diverse regarding AI readiness and infrastructure, additional resources, guidelines, and education would also help organizations assess decisions on investing in AI.[479, 480, 481]

HHS and other federal actions to date (non-exhaustive):

- **ASTP LEAP in Health Information Technology cooperative agreement awards** provided funding opportunities for the advanced development of AI solutions for patient care.[482]
- **CMS AI Health Outcomes Challenge** provided innovators an opportunity to showcase their AI tools that can be used to predict patient health outcomes for Medicare beneficiaries for potential use in CMS with an opportunity to showcase their AI tools that help predict patient health outcomes for Medicare beneficiaries, which could be used in CMS's innovative payment and service delivery models.[483]
- **NIH COVID-19 medical imaging** during the COVID-19 pandemic engaged in a multi-institutional effort utilizing medical imaging techniques screening for infected heart and lung features to assess disease severity and propose treatments.[484]
- **National Institute of Mental Health's (NIMH) Digital Global Mental Health Program** funds research on the development, testing, implementation, and cost-effectiveness of digital mental health technology appropriate for low- and middle-income countries.[485] It places emphasis on research leveraging AI and/or other novel computational and statistical approaches to improve the prevention, diagnosis, and treatment of mental health along a treatment trajectory and continuum of care.
- **General Service Administration's (GSA's) Technology Transformation Services (TTS) and other programs** support research in healthcare delivery, including through tech uplift and innovation support, and could be expanded to include AI.[486]
- **SAMHSA Innovative Uses of Technology to Enhance Access to Services Within the Crisis Continuum publication** highlights innovative uses of technology that help those in need get access to critical services, including how AI can help with disease screening and delivery (e.g., personalized self-serve mental health apps).
- **AHRQ AI and Healthcare Safety NOFO** invites grant applications that support healthcare safety by determining (1) whether and how certain breakthrough uses of AI systems can affect patient safety and (2) how AI systems can be safely implemented and used.[487]

[478] https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf

[479] https://pmc.ncbi.nlm.nih.gov/articles/PMC9628307/#

[480] https://pubmed.ncbi.nlm.nih.gov/30802901/

[481] https://www.nejm.org/doi/full/10.1056/NEJMra2204673

[482] https://www.hhs.gov/about/news/2024/09/17/hhs-announces-2024-leap-health-awardees-focused-data-quality-responsible-ai-accelerating-adoption-behavioral-health.html

[483] https://www.cms.gov/priorities/innovation/innovation-models/artificial-intelligence-health-outcomes-challenge

[484] https://www.nih.gov/news-events/news-releases/nih-harnesses-ai-covid-19-diagnosis-treatment-monitoring

[485] https://www.nimh.nih.gov/about/organization/cgmhr/digital-global-mental-health-program

[486] https://tts.gsa.gov/

[487] https://grants.nih.gov/grants/guide/pa-files/PA-24-261.html

- **SAMHSA Neural Network Analysis** utilizes an AI neural network to analyze the co-occurrence of substance use problems, anxiety disorders, and depressive orders.[488] Findings show evidence that mental health clinics should provide integrated treatment plans and screen for various conditions and factors.

HHS near-term priorities:

- Support health services research on best practices for procuring, deploying, and monitoring AI tools in healthcare delivery settings (e.g., **AHRQ healthcare safety and AI NOFO**).[489]
- Build on existing "challenge" initiatives driving innovation in AI relevant to healthcare delivery, such as the **CMS AI Health Outcomes Challenge** and the **NIH CRDC AI Data-Readiness (AIDR) Challenge**.[490, 491]
- Explore opportunities to expand initiatives that promote AI innovation in healthcare delivery contexts, such as the **GSA's TTS**.[492]
- Provide guidelines on how to test and pilot AI applications within healthcare institutions before fully implementing them in care delivery.

2. Providing guidelines and resources on oversight, medical liability, and privacy and security protections to increase confidence for organizations to develop and deploy AI

Context:

Providers are reticent to deploy new AI interventions without knowing whether they have been "vetted" by appropriate entities or whether these entities have considered patient outcomes, safety, privacy and other factors. They are further reluctant to use new AI technologies without appropriate clarity on their potential liability from using these tools.

First, on oversight of quality assurance and vetting of AI interventions, despite many regulations that address technology in healthcare (e.g., medical technologies including EHRs and RCM), there are still gaps in clarity and scope in how they may specifically address AI use (generally and situationally). For example, some AI technologies may fall outside of existing medical device authorities. Authority over the regulation of health IT that are not medical devices belongs in part to the ASTP/ONC. As described in the Medical Product Development, Safety, and Effectiveness chapter, **ASTP's HTI-1 Final Rule** does not create an approval process per se but does establish policies that require transparency on the part of certain certified health IT (such as EHRs) regarding the AI-based technology offered in such products. Starting on January 1, 2025, regulations finalized in the final rule require the availability of specific "source attribute" information for any decision support intervention technologies certified to 45 CFR 170.315(b)(11) (including AI-based decision support interventions) offered as part of the health IT product.[493] An increasing number of AI tools in health IT could fall outside of current regulation, including certain EHR-integrated AI decision support tools (e.g., appointment no-show prediction algorithms) and AI algorithms deployed by health plans and insurance issuers for utilization management and prior authorization. These tools that do not meet the statutory definition of "device" for FDA oversight may not currently undergo regulatory review, validation, or testing.[494] Additionally, the **HTI-1 Final Rule** applies to AI-based technologies regardless of device definitions, use cases (e.g., clinical, administrative), or risk categories. HHS aims to further refine its regulatory framework covering AI technologies to promote safe and trustworthy use.

---

488 https://www.tandfonline.com/doi/full/10.1080/15504263.2024.2357623

489 https://grants.nih.gov/grants/guide/pa-files/PA-24-261.html

490 https://www.cms.gov/priorities/innovation/innovation-models/artificial-intelligence-health-outcomes-challenge

491 https://commons.cancer.gov/news/nci-crdc-artificial-intelligence-data-readiness-aidr-challenge

492 https://tts.gsa.gov/

493 https://www.healthit.gov/topic/laws-regulation-and-policy/health-data-technology-and-interoperability-certification-program

494 https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software

Additionally, regarding liability, while there is considerable experience regarding liability associated with the uses of technology in medical practice, AI (and especially GenAI such as LLMs) "raise[s] distinctive issues that do not apply to older forms of CDS or ways of researching medical questions online."[495] The use of patient data in AI has caused concerns among both medical professionals and patients. These include how it can be used in model development, patient consent for providers and developers regarding data storage, and when patients are informed of use. Ongoing updates to model inputs and training make it difficult to establish fact patterns and/or recreate specific incidents or scenarios needed for evidentiary rules.

Regarding patient data usage, **HIPAA Privacy and Security Rule** compliance is required when covered entities or business associates use or disclose PHI for AI development or maintenance. Uses and disclosures of PHI under HIPAA require written patient authorization unless permitted for certain specified purposes such as treatment, payment, or healthcare operations. When PHI is used for research involving AI, depending on the type of PHI being disclosed and the type of research being conducted, the **HIPAA Privacy Rule** may require that the individual authorizes the use or disclosure of PHI or provide a waiver or alteration of authorization by an IRB.[496] Sharing PHI with AI developers may also create additional complexity. Ultimately, using or disclosing patient data, including PHI, for AI models requires case-specific assessment and management to ensure compliance with HIPAA and other privacy regulations.[497] The "Promote Trustworthy AI Development and Ethical and Responsible Use" section of this action plan further discusses patient security and privacy.

Ultimately, supplementing guidelines and regulations while enhancing clarity on oversight and quality assurance from HHS divisions will enhance confidence in adopting safe and appropriate AI use cases within delivery and financing.

HHS near-term priorities:

- Provide additional guidelines on how AI use in healthcare should adhere to **privacy and security standards, including HIPAA.** This will include providing guidelines on risks of re-identification in the context of HIPAA and delineating when data used for AI requires patient authorization (i.e., research). To execute this priority, HHS will collaborate with other federal agencies to create unified standards and frameworks for privacy and security in AI applications.
- Within applicable existing HHS and division authorities, provide additional guidelines on liability considerations for clinicians and healthcare providers using AI.
- Provide guidelines and frameworks for appropriate approaches and roles clinicians and support staff should have in engaging with AI (e.g., role suitability related to technology based on the risk level of the AI application).
- Continue to clarify and build stakeholder awareness on applicable oversight and regulatory structures.

3. Ensuring developers and potential deployers of AI have clarity on coverage and payment determination processes to encourage development of AI

Context:

With many providers already facing economic pressure, there is limited appetite to invest in or use new and emerging information technologies, particularly when there is no guarantee of payment for services.[498] Clear

[495] https://jamanetwork.com/journals/jama-health-forum/fullarticle/2805334

[496] 45 CFR 164.512(i)(1)(i)

[497] https://www.justice.gov/opcl/privacy-act-1974

[498] https://pmc.ncbi.nlm.nih.gov/articles/PMC8166111/#

frameworks for payment for AI-enabled services will influence the wider use of AI in medicine, as providers may be more financially incentivized to utilize such technologies.[499]

Increasing the clarity on frameworks for payment for AI services will require policymakers to disseminate information to technology developers, device manufacturers, clinicians, and patients. Clarity in payment determinations processes could support numerous priorities, including informing access to innovative technologies, reducing uncertainty for developers and manufacturers, protecting the safety of beneficiaries of federal programs, stewardship of federal funds, and encouraging evidence development where gaps exist.

HHS actions to date (non-exhaustive):

- **CMS established separate payment pathways for at least eight AI/ML-enabled devices** through CPT® and new technology add-on payments (NTAP) under the Medicare Inpatient Prospective Payment System (IPPS), as of May 2024,[500] which represents less than 5% of FDA-authorized AI-based products.[501, 502, 503] CMS has taken steps to ensure that Medicare coverage determination and payment pathways are clear for innovations, including those enabled by AI.
- **CMS payment for Software as a Service** (as referenced in the Medical Product Development, Safety, and Effectiveness chapter) established payment pathways for hospital outpatient departments through add-on codes.[504]
- **CMS Transitional Coverage for Emerging Technologies (TCET) (CMS-3421-FN)** (as referenced in the Medical Product Development, Safety, and Effectiveness chapter) finalized the TCET Pathway in August 2024 to facilitate safer and more predictable access to new technologies for Medicare beneficiaries and further reduce uncertainties about coverage.[505]

HHS near-term priorities:

- Convene key stakeholders to inform coverage process and requirements for federal insurance programs (e.g., policymakers, technology developers, device manufacturers, clinicians, and patients).
- Provide guidelines and clarity on the coverage determination process for new AI products and services provided to federal beneficiaries.
- Develop guidelines for AI developers regarding evidentiary standards for payment and coverage decision-making.

3.6.2 Promote Trustworthy AI Development and Ethical and Responsible Use

A primary focus of AI in care delivery is ensuring patient safety, security, and privacy. AHRQ defines patient safety as a multifaceted discipline intended to protect patients in care administration. Potential AI-related patient adverse events (resulting either from an incorrect action carried out by an AI-enabled tool or healthcare staff incorrectly using an AI-enabled tool) must be thoroughly mitigated. In healthcare delivery, some methods of increasing trustworthiness and safety related to AI include ensuring a human [is] in the loop during AI decision-making, ensuring that models and their use by providers and payers are transparent, interpretable, and explainable,

[499] https://doi.org/10.1038/s41746-022-00609-6 Paying for artificial intelligence in medicine

[500] https://www.nature.com/articles/s41746-022-00609-6/tables/1

[501] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

[502] https://doi.org/10.1038/s41746-022-00609-6 Paying for artificial intelligence in medicine

[503] As with other technologies, Medicare provides payment for AI-enabled devices on a case-by-case basis, based on applications submitted by healthcare providers, device manufacturers, or other stakeholders.

[504] https://www.govinfo.gov/content/pkg/FR-2022-11-23/pdf/2022-23918.pdf#

[505] https://www.cms.gov/newsroom/fact-sheets/final-notice-transitional-coverage-emerging-technologies-cms-3421-fn

and clear guardrails are established for its use. Lack of explainability in AI systems can lead to skepticism, over-reliance, or rejection by clinicians.[506, 507]

Underpinning these principles are the following priority areas where HHS can support the safe use of AI:

1. Enhancing enforcement and clarifying guidelines relating to existing requirements
2. Providing guidelines and support related to internal governance
3. Promoting external evaluation, monitoring, and transparency reporting
4. Enhancing infrastructure to ensure patient safety

Below, HHS discusses the context, its actions to date, and plans to promote trustworthy AI development and ethical and responsible use in healthcare delivery.

1. Enhancing enforcement and clarifying guidelines relating to existing requirements

Context:

As discussed earlier, HHS and its divisions can promote adoption and protect beneficiaries in the context of AI by clarifying existing healthcare regulations and proactively enforcing existing legal requirements that certain AI applications may violate. These efforts will provide an improved, safer patient experience in cases where questions exist about whether existing federal requirements are being properly applied.

Given the rapidly expanding nature of AI risks, adding clarity to existing regulations may not always be sufficient. HHS could develop new levers, rules, and programs to ensure that healthcare organizations and AI developers adhere to best-practice risk mitigation principles at every stage of the AI life cycle, spanning design, development, deployment, maintenance, and retirement.

HHS actions to date (non-exhaustive):

- **AHRQ AI developed a program in healthcare safety** (see subsections below for an additional dedicated discussion of patient safety) in response to EO 14110 and as part of the Patient Safety Organizations Program to allow for the rapid development of AI patient safety-focused data, analyses, and resources. The program helps collectively track and identify situations where AI deployed in healthcare settings may cause adverse events and provides a means of learning from such occurrences in the future. The Patient Safety and Quality Improvement Act of 2005, 42 U.S.C. 299b-21 et seq., which established the PSO Program, also provides certain legal protections for organizations to share information on patient safety events to improve care without the fear that the information could be used against them in settings such as legal or administrative proceedings.[508]
- **HHS Plan for Promoting Responsible Use of Artificial Intelligence in Automated and Algorithmic Systems by STLT Governments in the Administration of Public Benefits** includes recommendations such as impact assessment to determine estimated benefits and risks from AI, measuring the quality and appropriateness of the data used in a system's training, testing, and prediction, and consulting workers and providing adequate training for all staff around developing, using, enhancing, and maintaining automated and algorithmic systems.[509]
- **Final Rule on Nondiscrimination in Health Programs and Activities (**Section 1557 of the Patient Protection and Affordable Care Act ["Section 1557"]) prohibits discrimination in certain health programs and activities, and, like other federal civil rights laws, Section 1557 applies to the use of AI, clinical algorithms, predictive analytics, and other tools.

[506] https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01332-6

[507] https://ccforum.biomedcentral.com/articles/10.1186/s13054-024-05005-y#

[508] https://pso.ahrq.gov/sites/default/files/wysiwyg/ai-healthcare-safety-program.pdf

[509] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

- Section 1557 includes a provision that applies non-discrimination principles to using patient care decision support tools, including AI. It requires those organizations covered by the rule—including any health program or activity that receives Federal financial assistance from HHS, including health insurance exchanges and HHS health programs and activities—to take steps to identify and mitigate the risk of discrimination that may result through the use of AI and other forms of patient care decision support tools.[510, 511]

- **Frequently asked questions (FAQ) related to coverage criteria and utilization management requirements in CMS final rule (CMS-4201-F)** emphasized compliance with existing coverage rules by addressing the question of whether Medicare Advantage (MA) rules on coverage criteria prohibit MA organizations from using algorithms or AI to make coverage decisions. The FAQ response explained that while an algorithm may be used to assist in making coverage determinations, it is the responsibility of the MA organization to ensure compliance with applicable rules for coverage determinations, such as those related to medical necessity and basing a decision on individual patient's circumstances.
  - CMS released a rule on December 10, 2024, that provides additional clarifications on the topics of coverage criteria, utilization management requirements, and AI use that also clarifies and amends language in 422.112(a)(8) (Ensuring Equitable Access to Medicare Advantage (MA) Services—Guardrails for Artificial Intelligence).[512, 513]
- The **HHS Trustworthy AI playbook** details principles organizations can implement to foster additional trust in their AI development. It captures mandates from regulations such as EO 13960, Office of Management and Budget (OMB) Memorandum M-21-06, and NIST guidelines.[514]

HHS near-term priorities:

- Increase the oversight and enforcement of existing federal laws and regulations, such as those prohibiting denying medically necessary, covered services or discrimination in access to federal benefits.
- Collaborate with other agencies outside of HHS (e.g., the Federal Trade Commission [FTC]) to strengthen and enforce consumer protections related to health data privacy and false marketing in the context of AI. This could include monitoring and addressing AI applications that compromise health data privacy or enhancing data sharing among agencies to detect and respond more rapidly to AI violations in healthcare settings.
- Develop additional targeted guidelines building on existing policy frameworks that explain to regulated entities how to comply with existing requirements when AI tools or technologies are applied.

2. Providing guidelines and support related to the local governance of AI

Context:

Some healthcare delivery and financing organizations have established their governance frameworks and means of vetting, evaluating, and monitoring AI tools locally. However, in other cases, risk assessment

[510] See 45 Code of Federal Regulations (CFR) 92.210.

[511] These requirements will take effect on May 1, 2025.

[512] https://public-inspection.federalregister.gov/2024-27939.pdf

[513] https://www.ecfr.gov/current/title-42/chapter-IV/subchapter-B/part-422/subpart-C/section-422.112

[514] https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf

processes and clear governance structures may not be in place or may not be rigorous enough to protect patients or beneficiaries.[515]

HHS actions to date (non-exhaustive):

- **HHS and division playbooks** provided perspectives on risk, including the Trustworthy AI (TAI) playbook and CMS AI Playbook, which provide specific considerations to help organizations safely operationalize AI development.
- **AHRQ Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Healthcare** included principles for organizations seeking to mitigate racial and ethnic disparities across every step of the AI life cycle.[516]
- **AHRQ Digital Healthcare Equity Framework and Practical Guide for Implementation** is an evidence-based guide to help organizations intentionally consider equity in developing and using digital healthcare technologies and solutions. The Guide serves as a resource to digital healthcare developers and vendors, healthcare systems, clinical providers, and payers and includes a checklist of steps and real-world examples for advancing equity across phases of the Digital Healthcare Life Cycle.[517, 518]

HHS near-term priorities:

- Within HHS authorities, support efforts to develop targeted guidelines on risk management and internal AI governance for health organizations that build on existing policy, governance, and risk management frameworks (e.g., **NIST AI Risk Management Framework** and **ASTP's HTI-1 Final Rule**). Guidelines may include standards that apply globally, by sector, or by use type and are specific enough to apply effectively to different healthcare delivery and financing subcategories. They may vary by applicable division or framework.
- Explore using federal programs and incentives, including those administered by CMS, to require or encourage internal governance mechanisms and evaluation practices for healthcare delivery and financing organizations. This could include regulations requiring the establishment of internal committees responsible for monitoring and reviewing all AI use cases across their organizations.
- Explore mechanisms to ensure that healthcare delivery and financing organizations, including those administered by CMS, meet the minimum governance and evaluation standards and identify relevant authorities to enforce these requirements (e.g., via audits, corrective action plans, and enforcement in the event of continued noncompliance).
- Develop recommended minimum standards for evaluating the risk of AI tools. These could include risk stratification guidelines based on the device's potential impact and risk-appropriate monitoring cadence and metrics.
- Develop hospital guidelines and resources to identify, manage, and mitigate AI-related safety, bias, or effectiveness concerns.

HHS long-term priorities:

- Continue educating the public and clinical teams on trustworthy and safe AI through publications, research, and standards to interpret AI, communicate interventions to patients, identify types of adverse events that can occur with AI, and how to report such events through existing systems.
- Convene and/or support publicly accessible conferences and dialogue with industry experts on AI risks and appropriate risk management approaches.

---

[515] https://ai.nejm.org/doi/abs/10.1056/AIp2300048 For example, one study found that even well-resourced academic medical centers sometimes found it difficult to identify and manage potential problems associated with predictive AI tools. How Academic Medical Centers Govern AI Prediction Tools in the Context of Uncertainty and Evolving Regulation.

[516] https://pubmed.ncbi.nlm.nih.gov/38100101/

[517] https://digital.ahrq.gov/health-it-tools-and-resources/digital-healthcare-equity

[518] https://digital.ahrq.gov/health-it-tools-and-resources/digital-healthcare-equity/digital-healthcare-equity-framework-and-guide

3. Promoting external evaluation, monitoring, and transparency reporting to enhance the quality assurance of AI

Context:

Testing and evaluating the effects of AI in real-world delivery settings is challenging due to the rapid expansion of AI in different clinical areas and due to common challenges inherent to clinical medical data (e.g., low prevalence of certain diseases, lack of or difficulty in obtaining ground truth data). Furthermore, the potential of AI lies in its ability to design models that learn, update, and adapt continuously as more data becomes available or as data changes. This ability poses unique regulatory challenges, requiring the development of suitable controls and testing methods that balance the potential benefits and risks of adopting AI within and beyond traditional clinic settings. HHS recognizes these real-world challenges associated with establishing appropriate evaluation and monitoring processes and will balance the scope of required monitoring and evaluation against the risk posed by AI in proposing regulatory guardrails.

Given the large volume and diversity of anticipated AI applications needing some evaluation and the need to take local considerations into account, HHS anticipates the need for a public/private approach to quality assurance of AI used in healthcare.[519] To help anchor a nationwide quality assurance approach, HHS may consider whether there are areas where rulemaking may be appropriate to enable successful governance practices and oversight of the use of AI in healthcare delivery and financing, for example, by motivating and supporting nationwide public-private approaches to validate AI. *See also the Medical Product Development, Safety, and Effectiveness chapter for further discussion of approaches to quality assurance of health AI.*

HHS actions to date (non-exhaustive):

- **HTI-1 Final Rule** lays a foundation for transparency by establishing a set of requirements for certain AI supplied by EHR developers and their systems that are certified under the ONC Health IT Certification Program to ensure clinical users will be able to access a consistent, baseline set of information about the algorithms they use to support their decision-making.[520]

Other industry actions to date (non-exhaustive):

- Multiple organizations collaborating on initiatives to convene healthcare delivery stakeholders to address challenges and launch initiatives related to AI.[521]
- **The Trustworthy and Responsible AI Network (TRAIN)** is a collaboration of provider organizations working to operationalize responsible AI principles.[522]
- **The Coalition for Health AI (CHAI)** is a collaboration among healthcare organizations and technology developers to promote development, evaluation, and appropriate use of AI. The collaboration has developed a template "model card" aligned with ASTP's HTI-1 AI transparency requirements.[523]

HHS near-term priorities:

- **Build on transparency requirements** by working with the industry to specify consensus approaches to standardized metrics, information, and data for HTI-1's decision support interventions' source attribute (aka "model card") requirements.

---

[519] https://jamanetwork.com/journals/jama/article-abstract/2813425 A Nationwide Network of Health AI Assurance Laboratories.

[520] Providers and payers have voluntarily committed to leveraging this framework to help guide their AI governance, development, and purchasing activities.

[521] https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000513

[522] https://train4health.ai/

[523] https://chai.org/draft-chai-applied-model-card/

- **Support efforts to widen the accessibility of AI performance information** by considering incentives to disclose healthcare providers' and payers' use of AI and related performance information that impacts access to or the quality of care.
- **Explore the use of federal programs and incentives** to encourage external accountability mechanisms for payer and provider organizations deploying AI, including:
    - Motivating deployers of AI to undergo independent, external algorithmic audits conducted by certified entities free from conflict of interest
    - Incentivizing performance transparency among other developers and deployers of AI to include a broader range of technologies (e.g., beyond EHR technologies covered by policies finalized in HTI-1)[524]
    - Collaboration with existing and emerging validation, monitoring, and transparency efforts in the private sector, supporting when and where appropriate.

4. Enhancing infrastructure to ensure patient safety, security, and privacy

Context:

As discussed previously in this Plan, maintaining patient safety, security, and privacy is a pivotal but complex challenge compounded by the possibilities of AI to influence or administer care delivery. A key component of ensuring patient safety is maintaining enough direct oversight by clinical staff (including face-to-face time between doctors and patients and monitoring of insights that AI may suggest). Patients also increasingly demand transparency about decisions impacting their care, particularly if AI tools influence diagnoses or treatments. Caregivers also require clear information on their AI tools so they can communicate to patients how AI is being leveraged in care, which will empower patients to make informed decisions and provide consent. AI introduces potential new vulnerabilities concerning patient security and privacy. The types of data demanded and the number of stakeholders seeking it continue broadening, underscoring the importance of ensuring robust patient data protection as AI use expands. As discussed in "Catalyze health AI innovation and adoption" above, HHS privacy and security protections such as HIPAA provide guidelines for handling patient data. Still, an additional opportunity exists to evolve such protections in parallel with AI technology. HHS has already taken steps to address such areas of concern for patient safety, security, and privacy in AI and will continue expanding its strategy.

HHS actions to date (non-exhaustive):

- **AHRQ's AI in Healthcare Safety Program** took steps to analyze and aggregate data of types of AI incidents (e.g., patients, caregivers, or others) and encourages more organizations to work with PSOs that support patient safety and quality improvement. Specifically, to "establish a common framework for approaches to identifying and capturing clinical errors resulting from AI deployed in healthcare settings," the existing common formats provide a basis for and can be enhanced to better capture such concerns.[525]
- **AHRQ's PSO Program—communication mechanism—**engaged with the PSOs on AI and healthcare through various presentations and discussions.[526]
- **AHRQ exploratory analyses of patient safety events,** through the Network of Patient Safety Databases (NPSD), analyzed potential AI-related patient safety events to better understand the current

[524] Metrics and measures that are similar to what pertain to certified EHRs would provide users basic information about algorithms, training data, and performance metrics and provide a better foundation for evaluation.

[525] https://pso.ahrq.gov/resources/ai-healthcare-safety#

[526] https://www.psoppc.org/psoppc_web/DLMS/downloadDocument?groupId=2371&pageName=welcome

capacity of the Common Formats and NPSD in capturing where AI deployed in the healthcare setting may cause unintended impacts.[527, 528]

HHS near-term priorities:

- Expand the capability of PSOs to assist providers in learning from and preventing potentially AI-related adverse impacts through education, resource sharing, and development.
- Explore mechanisms to encourage data submission on potentially AI-related events as part of the **AI in Healthcare Safety Program.**

HHS long-term priorities:

- Utilize the **NPSD** as the "central tracking repository" for patient safety incidents resulting from AI deployed in healthcare settings. The repository already includes some related information.
- Consider expanding **AHRQ AI in Healthcare Safety Program** to sustain and build upon initial program projects and advance activities that analyze the NPSD.[529]
- Promote AHRQ research on mitigating racial bias from algorithms[530]
- Consider expanding grant-making projects and NOFOs.
- Continue to evaluate potential impacts that AI may have on patient-provider interactions (e.g., direct face-to-face time, gathering of patient information).

3.6.3 Democratize AI Technologies and Resources

To achieve the goals for AI to accelerate access and equity in healthcare delivery, the technology and understanding around implementation must be accessible. Without the explicit consideration of biases resulting from the under-representation of certain patient populations from training data, underserved settings could find themselves experiencing less benefit from AI.[531]

HHS can implement actions in the areas below, with particular attention to stakeholder groups that may already be affected by the digital divide:

1. Promoting equitable access through technical support for and collaboration with delivery organizations that provide services to underserved populations
2. Providing support for healthcare delivery organizations to address core infrastructure and deployment challenges (i.e., technology, infrastructure, and data infrastructure) that improve AI readiness

Below, HHS discusses the context, its actions to date, and plans to democratize AI technologies and resources within the healthcare sector.

1. **Promoting equitable access through technical support for and collaboration with delivery organizations that provide services to underserved populations**

   Context:

   The extent of possible AI impacts on underserved populations is still greatly unknown, especially given the complex and under-researched nuances that underserved communities may face.[532]

[527] https://www.psoppc.org/psoppc_web/DLMS/downloadDocument?groupId=2372&pageName=welcome
[528] https://pso.ahrq.gov/common-formats
[529] https://pso.ahrq.gov/resources/ai-healthcare-safety#
[530] https://pubmed.ncbi.nlm.nih.gov/38100101/
[531] https://pmc.ncbi.nlm.nih.gov/articles/PMC10844447/
[532] https://pmc.ncbi.nlm.nih.gov/articles/PMC8486995/

Care providers in predominately underserved settings—e.g., community health centers and safety-net hospitals—may stand to benefit the most from the potential of AI (e.g., reduced costs, lower administrative burdens) while also facing the largest barriers, given a lack of AI expertise and robust capital budgets to deploy new technology.

As discussed earlier in the "Trends" section of this chapter, additional concerns about equitable delivery of care come from the potential of AI to automate traditional interactions administered by providers. For populations in underserved settings, the reduction of patient-provider interactions poses risks to patients, especially when social factors such as literacy and culture directly impact patient experience. Additionally, there is an increased risk for such populations if care settings become less appropriately staffed because of AI. HHS is committed to helping organizations determine which technologies are most suitable for their contexts and collaborating with underserved populations to increase research efforts on how AI can impact care delivery and outcomes.[533, 534]

HHS actions to date (non-exhaustive):

- **NIH's AIM-AHEAD Program** increases diversity in AI researchers and data by providing underrepresented communities with AI access through partnerships, research, infrastructure, and data science training to expand the participation and representation of currently underrepresented populations in developing AI models.[535]

HHS near-term priorities:

- Establish regional technical assistance centers through grants or cooperative agreements that can aid under-resourced care settings on AI applications.
- Disseminate AI impact assessment templates, implementation toolkits, and technical assistance resources for health delivery organizations considering using AI by either promoting existing tools or funding the creation of new tools where gaps exist.
- Fund research to develop insights on best practices for adopting AI applications in under-resourced settings. This may include helping under-resourced organizations run pilots of high-potential AI.

HHS long-term priorities:

- Convene communities of practice across healthcare delivery to facilitate information sharing on the application of AI, particularly in underserved populations. This may include soliciting feedback and input from organizations in underserved populations that have adopted AI (e.g., through already available assistance from a private or non-profit entity) on addressing key challenges.
- Continue to evaluate potential impacts that AI may have on patient-provider interactions (e.g., direct face-to-face time, gathering of patient information).

2. Providing support for healthcare delivery organizations to address core infrastructure and deployment challenges (i.e., technology, infrastructure, and data infrastructure) that improve AI readiness

Context:

Organizations need prerequisite capabilities and infrastructure including data systems to leverage AI. Healthcare delivery is a vastly complex system with various specialties and stakeholders administering care. As such, infrastructural tools and AI will not likely be generalizable to broad types of hospitals and clinical settings. For example, pediatric specialties face a distinct set of circumstances compared to adult specialties,

[533] https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/006_Serving_Vulnerable_and_Underserved_Populations.pdf
[534] https://www.politico.com/newsletters/future-pulse/2024/04/25/ai-degrades-our-work-nurses-say-00154253
[535] https://datascience.nih.gov/artificial-intelligence/aim-ahead

underscoring the importance of children's hospitals and pediatric units having the flexibility to configure AI to their contexts.

While data quality and accuracy are necessary for training algorithms, the availability of datasets for training and tuning is an industrywide barrier to developing higher-quality health AI, especially for smaller and under-resourced healthcare delivery and payer organizations. Additionally, the technological infrastructure to sufficiently run AI models and store large data is not yet widely accessible to or affordable by healthcare organizations, limiting their ability to utilize AI at scale.

Hospitals and ambulatory practices that benefit from federal incentives to adopt technology such as EHRs may be better placed to adopt AI because of the availability of AI tools and third-party vendors integrating through EHR. Providers with a lower adoption of EHR technology, such as behavioral health and long-term post-acute care entities, may find AI tools less available and usable.

HHS actions to date (non-exhaustive):

- **HRSA Uniform Data System (UDS) Modernization Initiative** updated and improved the UDS dataset and the technology used for data submission, collection, and analysis, providing HRSA with de-identified patient data. This initiative enables HRSA to support its nearly 1,400 health centers better by implementing more effective data analysis and predictive solutions.[536]
- **IHS Data Modernization Initiative** partners with tribal and urban leaders to modernize EHR standards across the IHS system to enhance interoperability and functionality in healthcare to serve patients better.[537]
- **ACF Data Strategy** increased its AI capabilities and supported analysis on care delivery opportunities for children, families, and underserved populations through increased data interoperability.[538, 539]

HHS near-term priorities:

- Work with the industry to promote open-source AI specifications for stakeholders (e.g., developers) to leverage.
- Establish regional technical assistance centers through grants or cooperative agreements to support lower-resourced care settings—specifically on data and technology modernization to enhance AI readiness.
- Disseminate AI-readiness assessment templates for providers considering developing AI solutions to support decision-making on data, system, and technology infrastructure gaps.

HHS long-term priorities:

- Update internal data infrastructure to ensure sufficient and actionable information is available for underserved communities to inform support strategies HHS can implement.
- Make available open de-identified data assets of administrative, clinical, quality/outcomes, and safety data to support AI development, testing, and validation.

HHS has established similar resources for providers seeking to implement EHRs or undertake quality improvement or payment reform initiatives. The department can leverage the experience of implementing those initiatives, but it would likely require additional funding to establish these additional AI deployment resources.

[536] https://bphc.hrsa.gov/data-reporting/uds-training-and-technical-assistance/uniform-data-system-uds-modernization-initiative#

[537] https://www.ihs.gov/hit/

[538] https://www.acf.hhs.gov/ai-data-research/artificial-intelligence-acf

[539] https://www.acf.hhs.gov/ai-data-research/acf-data-strategy

3.6.4 Cultivate AI-Empowered Workforces and Organization Cultures

A workforce that is knowledgeable on AI will help accelerate innovation (e.g., identifying new use cases), manage deployment-specific risks associated with new tools, establish appropriate organizational governance structures, evaluate setting-specific training data for potential biases, monitor model-drift,[540] mitigate adverse impacts, and communicate with patients, families, and providers about the use of this technology. An effective health AI workforce will require cross-functional teams, including clinicians, biostatisticians, privacy/information security officials, analysts, acquisition staff, and IT professionals. Ensuring that individuals and organizations are sufficiently prepared to use AI will be critical in safe, effective, and widespread adoption.

The key opportunity HHS will focus on is **equipping healthcare delivery professionals with access to training, resources, and research to support AI literacy and expertise in their respective health system organizations.**

Context:

To date, AI is incorporated into the curriculums of most healthcare education, certification, and continuing education programs in a limited capacity, if at all. Additionally, most AI expertise is concentrated within technology organizations and/or research institutions (e.g., universities and large technology organizations). HHS will take steps to increase AI knowledge and expertise among healthcare professionals, ensuring foundational know-how within delivery organizations that lowers the cost of implementing new tools and ensures they are applied appropriately.

HHS actions to date (non-exhaustive):

- **CMS AI Playbook** included educational materials that define AI, use cases, and trends within healthcare delivery, along with applications that CMS is considering using within its operations and their potential impact on patient care (e.g., wearables, digital twins, customer support).[541]
- **AHRQ's intramural research programs** (e.g., Health Services Research Dissertation Awards, Institutional Training Awards, Mentored Clinical Scientist Development Awards) offered predoctoral and post-doctoral educational, research infrastructure, and career development grants and opportunities in health services research. In addition, the AHRQ supports the development of health services research infrastructure in emerging centers of excellence and works with Federal and academic partners to develop innovative curricula and educational models.[542]

[540] Monitoring model drift is essential to ensuring AI models and resulting diagnostic and therapeutic decisions are based on relevant data.

[541] https://ai.cms.gov/assets/CMS_AI_Playbook.pdf

[542] https://www.ahrq.gov/funding/training-grants/rsrchtng.html

HHS long-term priorities:

- Promote AI literacy through long-term public education initiatives focused on reaching an audience of professionals (clinical and non-clinical) operating in a healthcare delivery context
- Convene regular AI sessions at healthcare conferences with seminars hosted by industry experts, learning tracks, practical workshops, and recorded resources to promote collaboration, learning, and innovation.
- Develop guidelines on appropriate training curricula and cadence for how AI concepts should be covered across cadres of healthcare workers (e.g., continuing medical education, degree programs).
- Directly fund workforce training programs that train the existing health AI workforce and educate the next generation of medical professionals.
- Share AI internal training resources on public websites for health AI professionals working in the industry to adapt or use directly in various healthcare settings.
- Continue to evaluate the potential impacts of AI on the healthcare workforce.

The interventions listed above are focused primarily on developing new programs and using public education and outreach, including to varied populations such as people with disabilities, to promote the responsible use of AI in healthcare. Looking ahead, agencies should also review existing workforce training programs and funding sources for health services research that can be leveraged to accomplish these objectives.

3.7 Conclusion

Through actions in its Strategic Plan, HHS will help facilitate delivery organizations' ability to expand access and transform patient care using AI. Given the rapid advancements in AI, HHS will continually review the actions of this plan and make efforts to extend support to stakeholders in the healthcare delivery ecosystem.

4 Human Services Delivery

4.1 Introduction and Context

AI presents an opportunity to improve the quality, accessibility, interoperability, coordination, and overall impact of human services programs in the U.S. The aging and diversifying population, complex and disparate public benefits systems, and persistent workforce shortage heighten the potential of AI in the sector. However, AI adoption in human services is nascent, reflecting critical challenges, including a lack of funding, outdated IT and data infrastructure, and concerns over technology's impact on human services program participants.[543, 544]

Despite these challenges, interest in AI remains, with 83% of government leaders believing technology will become more important in supporting the human services workforce.[545] HHS has released its **Plan for Promoting the Responsible Use of Artificial Intelligence in Automated and Algorithmic Systems by State, Local, Tribal, and Territorial Governments in Public Benefit Administration.**[546] However, there is an opportunity to do more. HHS aspires to maximize the opportunities of AI while protecting Americans' safety and security by ensuring the technology is tested, deployed, and monitored responsibly.[547] HHS has identified actions as part of this Plan to catalyze AI innovation and adoption, promote trustworthy AI development and ethical and responsible use, democratize AI technologies and resources, and cultivate AI-empowered workforces and organization cultures.

In this chapter, HHS outlines the scope and stakeholders relevant to AI in human services delivery before providing an overview of the opportunities of AI in the sector and observed trends. The chapter then outlines potential use cases for AI in human services and the risks of AI adoption. Finally, it concludes with a proposed approach to meet HHS's departmentwide goals for AI, which considers gaps, existing initiatives, and new opportunities.

[543] https://nff.org/learn/survey 2022 survey of non-profits from the Nonprofit Finance Fund found that more than half of participating organizations felt they would be unable to meet demand for their services in the upcoming year.

[544] https://pmc.ncbi.nlm.nih.gov/articles/PMC6816239/ Review of funding models for evidence-based interventions. "Every traditional pot of funding has a little bit of a question mark on it."

[545] https://www.cpsai.org/. Cited on the home page from a survey conducted by the Center for Public Sector AI.

[546] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf Published in April 2024 (herein referred to as Plan for Promoting Responsible Use of AI in Public Benefits)

[547] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf Published in April 2024 (herein referred to as Plan for Promoting Responsible Use of AI in Public Benefits)

4.1.1 Scope of the Human Services Delivery AI Value Chain

Exhibit 92: Overview of HHS Human Services Delivery Programmatic Areas

	Promote health and well-being	Assist populations with complex needs	Support families and children	Enhance community and economic development
Description	Access to healthcare services, preventative care, treatment for illnesses, mental health support Public health initiatives to prevent disease and promote a healthy lifestyle	Aid to individuals experiencing economic hardship, homelessness, substance abuse Support for seniors, including long-term care and community services	Childcare, early care and education, family support Child welfare services, foster care, adoption	Community initiatives improving local infrastructure and services Economic assistance and job training

Note: other chapters including Public Health and Healthcare Delivery cover similar or interconnected services. However, as much as possible HHS has segmented discussion of human services programs into this chapter.

4.1.2 Action Plan Summary

Later in this chapter, HHS articulates proposed actions to advance its four goals for the responsible use of AI in the sector. Below is a summary of the themes of actions within each goal. For full details of proposed actions please see section 4.6 Action Plan.

Key goals that actions support	Themes of proposed actions *(not exhaustive, see 4.6 Action Plan for more details)*
1. Catalyzing health AI innovation and adoption	• Unlocking resources for AI adoption and modernizing IT and tech infrastructure • Ensuring data quality and availability for AI adoption
2. Promoting trustworthy AI development and ethical and responsible use	• Providing guidance to served populations on balancing risks with opportunities for AI applications and establishing participant trust
3. Democratizing AI technologies and resources	• Raising the floor of constituent digital literacy and digital penetration • Identifying areas of cooperation across sectors to improve AI-related economies of scale
4. Cultivating AI-empowered workforces and organization cultures	• Improving human services employee digital literacy, talent, and openness to adopt new technology • Using AI to mitigate the labor workforce shortage in human services

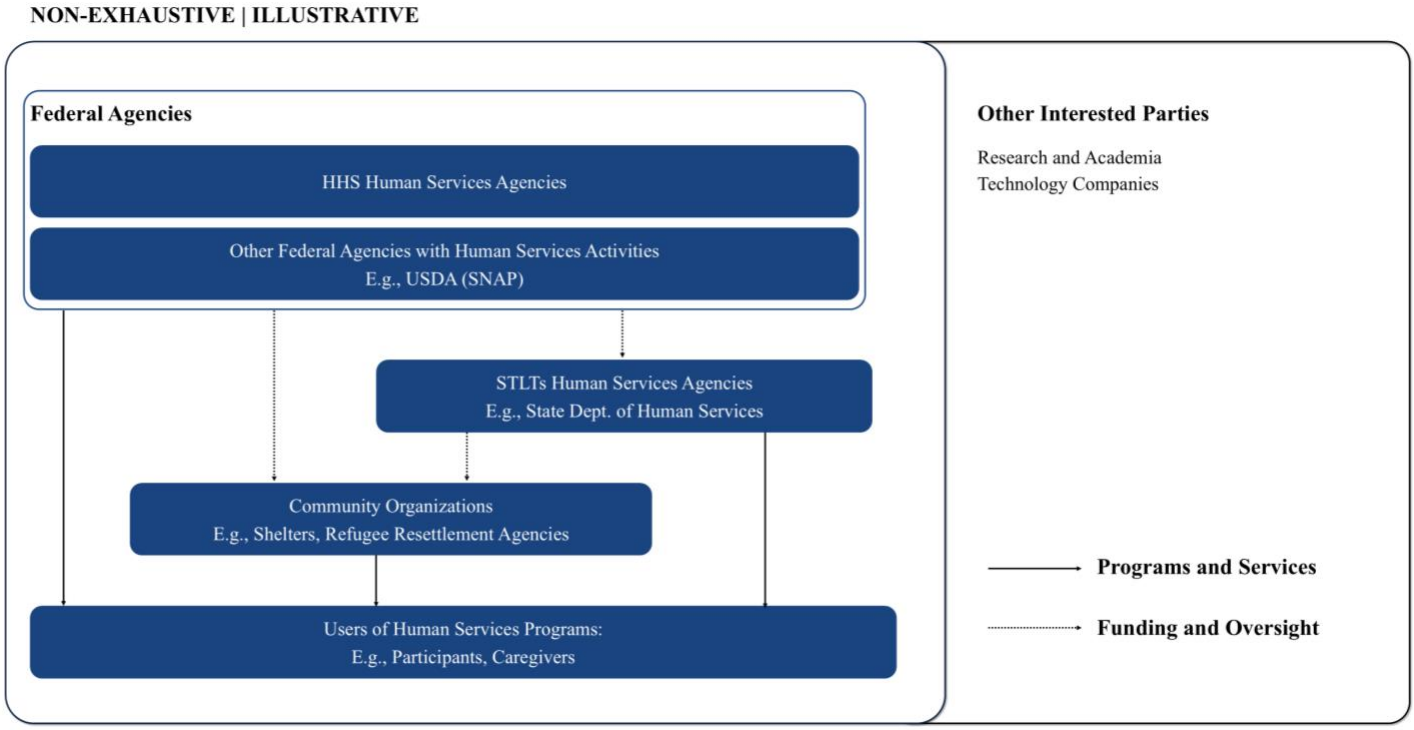
4.2 Stakeholders Engaged in the Human Services Delivery AI Value Chain

Human services programs in the U.S. benefit the most vulnerable populations, their caregivers, and their guardians. Various federal, STLT, and community stakeholders contribute to programs that serve that aim. Federal agencies fund STLT human services agencies and community organizations to deliver programs while also delivering programs themselves; STLTs fund community organizations and directly deliver programs; and CBOs deliver programs with a combination of federal, STLT, and philanthropic funds.

Exhibit 10 shows a non-exhaustive diagram of example flows between stakeholders and a bulleted list of stakeholders involved in human services. Please note that neither the diagram nor the list captures all roles and interactions. For additional details on regulatory guidance and authorities, please refer to other HHS documents.

The exhibit reflects example roles and relationships, but roles may vary depending on the human services program.

Exhibit 10: Human Services Delivery Stakeholder Engagement Map



- **HHS agencies**[548]
  - **ACF:** Provides services to support families and children, including promoting the economic and social well-being of children, families, and communities.
  - **ACL:** Supports programs for populations with complex needs, particularly older adults and people with disabilities.
  - **CMS:** Administers federal health insurance programs (e.g., Medicare and Medicaid), outlines conditions of participation related to these programs, and can provide reimbursements to specific devices or services.
  - **SAMHSA:** Focuses on promoting health and well-being, including services related to suicide prevention and mental health and substance abuse treatment and prevention.
  - **HRSA:** Provides access to essential health services for underserved populations, focusing on services that promote health and well-being and assist populations with complex needs.
  - **IHS:** Provides a comprehensive healthcare delivery system and ensures culturally appropriate public health and human services are available for American Indian and Alaska Native people to raise the physical, mental, social, and spiritual health of the population to the highest level.
- **Other federal agencies:** HHS also works closely with many other federal departments, such as the Department of Agriculture and the Department of Housing and Urban Development.
- **STLT government human services agencies:** STLT human services departments administer programs and provide public benefits. These departments often administer federal programs like the Supplemental Nutrition Assistance Program (SNAP) and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) alongside HHS programs.

[548] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf See Appendix B in the Plan for Responsible Use of AI for an overview of major human services and other public benefits programs administered by HHS.

- **CBOs, including community action agencies:** These organizations directly deliver human services programs and benefits to the public.
- **Participants and their caregivers and guardians:** In 2023, an estimated 99.1 million people (30% of the U.S. population) accessed services from various programs, including human services, collectively known as the "social safety net."[549] This figure includes one in eight adults and one in two children.
- **Technology companies:** These include companies focused on AI infrastructure (e.g., cloud storage), large, diversified tech companies, vendors of digital solutions, and white hat hackers. These companies provide the infrastructure and services for stakeholders to adopt AI.
- **Research institutions:** Often in partnership with federal agencies, STLTs, or CBOs, academic or other research institutions conduct trials and evaluations to understand the evidence for human services interventions and design and test potential programs.

4.3 Opportunities for the Application of AI in Human Services Delivery

AI in human services can improve service experience and quality, increase the pace and quality of funds distribution, enhancing capabilities of the human services workforce, increase accessibility of services, and enhance interoperability to improve service coordination. These opportunities are driven by multiple factors, including changing population demographics, a complex public benefits ecosystem, and workforce shortages.

The opportunities include:

1. **Improving service experience and quality:** Eligible participants face challenges accessing human services programs and consider the experience difficult.[550] Public sector health and human services have lower customer satisfaction scores than other industries surveyed by the American Customer Satisfaction Index.[551] AI can address challenges and improve satisfaction, including by assisting in matching participants to programs, speeding up application processes, improving benefit delivery speed, and enhancing the participant support experience.
2. **Increasing the pace and quality of funds distribution:** Billions of dollars flow through HHS to STLTs, community organizations, and directly to Americans around the country.[552] Often, the faster these funds can be appropriately distributed, the faster public benefits and vital services can be delivered.[553] As the assistance programs launched during COVID-19 pandemic demonstrated, the ability to quickly and effectively provide funds has the potential to save lives and livelihoods.[554] AI has the potential to improve the speed and accuracy of funding distribution from HHS to other stakeholders and ensure that resource distribution is equitable and linked to areas with the greatest need.
3. **Enhancing capabilities of human services workforce:** Human services departments face challenges in recruiting and retaining critical workforce populations, and needs are only growing. For instance, the Bureau of Labor Statistics projects an annual social worker shortage across the U.S. of 67,300 over the next decade.[555] Other estimates suggest the gap is closer to 100,000.[556] The workforce shortage can lead to longer wait times and reduced services.[557] At the same time, the American public is aging,[558] and more people are

[549] https://aspe.hhs.gov/sites/default/files/documents/18eff5e45b2be85fb4c350176bca5c28/how-many-people-social-safety-net.pdf

[550] https://www.urban.org/research/publication/customer-service-experiences-and-enrollment-difficulties. Difficulties included trouble determining eligibility, providing documentation, navigating varied requirements, and receiving benefits when needed.

[551] https://theacsi.org/news-and-resources/reports/2024/10/15/acsi-insurance-and-mortgage-lenders-study-2024/ A full industry comparison is available in the report.

[552] https://www.hhs.gov/sites/default/files/fy-2024-budget-in-brief.pdf Multiple examples including the HRSA Health Center Program that proposed awarding $7.1B to 1,400 health centers in 2024 or TANF which passed $17.3B in funding to states in FY 2023

[553] https://pmc.ncbi.nlm.nih.gov/articles/PMC6816239/

[554] https://www.cbpp.org/research/poverty-and-inequality/robust-covid-relief-achieved-historic-gains-against-poverty-and-0

[555] https://www.bls.gov/ooh/community-and-social-service/social-workers.htm

[556] https://www.cpsai.org/

[557] https://www.councilofnonprofits.org/nonprofit-workforce-shortage-crisis

[558] https://www.census.gov/newsroom/press-releases/2023/population-estimates-characteristics.html

expected to access human services programs over time.[559] This may strain workforce capacity, increase demand for services, and place greater emphasis on efficient benefits provisioning. AI can augment the human services workforce's processes by automating rote tasks, processing narrative information (e.g., client notes, meetings, interviews) with NLP, and drafting documents. In one analogous setting, customer support centers, a National Bureau of Economic Research study found that using an AI-based conversational assistant improved worker productivity by 14%.[560] An equivalent productivity enhancement in human services could allow staff to allocate more time to value-added tasks and participant interaction, increasing worker productivity even if staff shortages persist. HHS also acknowledges concerns related to potential staff displacement and outlines actions below to monitor workforce and service impacts.

4. **Increasing accessibility of services:** A diverse and growing population qualifies for human services, yet many struggle to access these programs. According to the Urban Institute, four in ten adults reported enrollment difficulties in accessing public services, including Temporary Assistance for Needy Families (TANF) and SNAP.[561] Multiple factors may drive enrollment accessibility challenges. For instance, benefits applications require advanced vocabulary, health literacy, or financial literacy.[562] However, according to the American Community Survey, 26 million U.S. residents (approximately 9% of the population) have limited English proficiency.[563, 564] Program access challenges can prevent eligible participants from accessing services when they are needed. For instance, during the COVID-19 pandemic, participation in WIC only grew by 2% from 2020 to 2021 despite an increase in eligibility.[565] AI can assist stakeholders in the human services delivery ecosystem by increasing access to their services and meeting their equity goals. AI applications have improved accessibility in other sectors through technologies like visual assistance and closed captioning. Further, advances in GenAI have improved the accuracy and cultural nuances of automated language translation.[566] Human services staff may not be able to solely rely on these tools, but they can adapt translation models to target participant needs and reach communities chronically underserved due to language gaps.[567, 568]
5. **Enhancing interoperability to improve service coordination:** A significant volume of human-services-related data is collected in narrative format (e.g., case notes) or manually transcribed (e.g., in shelters).[569] These records are not easily searchable and require manual review, hindering the data quality for accurate needs assessment, service delivery and care, systemwide analytics, or interoperability between agencies.[570] Another challenge for interoperability is the complex and multifaceted nature of the U.S. public benefits system. Those wishing to access programs must comply with varied administrative and program requirements to apply for services; however, these requirements are inconsistent across states and systems

[559] https://www.healthsystemtracker.org/chart-collection/health-expenditures-vary-across-population/, https://www.cdc.gov/pcd/issues/2024/23_0267.htm Extrapolated from healthcare spend and chronic disease trends in U.S. population

[560] https://www.nber.org/papers/w31161 As measured in issues resolved per hour

[561] https://www.urban.org/research/publication/customer-service-experiences-and-enrollment-difficulties

[562] https://www.cbpp.org/sites/default/files/11-18-08fa.pdf SNAP applications require an understanding of gross versus net income, and which assets count against eligibility (savings accounts) and which do not (property).

[563] https://www.census.gov/newsroom/press-releases/2017/acs-5yr.html

[564] https://www.kff.org/racial-equity-and-health-policy/issue-brief/five-key-facts-about-immigrants-with-limited-english-proficiency/

[565] https://www.cbpp.org/research/food-assistance/eligible-low-income-children-missing-out-on-crucial-wic-benefits-during

[566] https://www.sciencedirect.com/science/article/pii/S2772941924000012

[567] https://lfaidata.foundation/blog/2024/05/21/translation-augmented-generation-breaking-language-barriers-in-llm-ecosystem/ LLM projects to improve translation from English to less widely spoken languages

[568] https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-92/subpart-C/section-92.201 Section 1557 requires that, if a covered entity uses machine translation, the translation must be reviewed by a qualified human translator when the underlying text is critical to the rights, benefits, or meaningful access to an individual with limited English proficiency, when accuracy is essential, or when the source documents or materials contained complex, non-literal, or technical language.

[569] https://controller.lacity.gov/landings/interim-housing-audit The Los Angeles Homeless Services Authority (LAHSA) released an audit in 2023 that found inaccuracies in its shelter capacity data. The system is maintained with an email-based daily census report system which is centrally copied into a master file by hand.

[570] https://controller.lacity.gov/landings/interim-housing-audit The daily census reports did not meet the accuracy requirements for use by LAHSA's bed availability system.

vary.[571, 572] Advances in AI technologies, including optical character recognition, NLP, and LLMs, have the potential to transform data into structured formats and more easily improve service delivery and share them across agencies.[573] Further, higher-quality data could allow AI applications such as integrated benefits systems to shift delivery to a person-centered design, where benefits across healthcare, housing, family assistance, and food security are coordinated and delivered together.[574]

4.4 Trends in AI in Human Services Delivery

Current trends indicate that AI in human services is nascent, but interest in piloting innovative technology is growing among STLTs and community organizations:

1. **STLT and community groups are interested in AI adoption. Still, they are early in the process with a focus on ideation and collaboration:** Multiple non-profit organizations have established AI practices, indicating enthusiasm throughout the domain. Examples of actions in the human services ecosystem are articulated below; however, these are non-exhaustive:[575]
    a. *U.S. Digital Response* launched tools to help state and local governments safely use GenAI to do their jobs better and faster.[576]
    b. *Center for Public Sector AI* launched Rolling Prompts that allow companies and organizations working with AI to share their ideas with state health and human services leaders on how to apply their technology in the domain.[577]
    c. *GovAI Coalition* developed policy templates and other resources to support STLTs with implementing governance for responsible experimentation and use of AI. The coalition represents more than five hundred agencies, primarily from city and local governments.[578]
2. **However, AI adoption in the human services sector remains low despite the opportunities it presents:** The current adoption of AI at scale—beyond the pilot phase—is low in the human services ecosystem compared to other sectors.[579] While some private sector and non-profit players have launched programs, these are often limited in scope or targeted to a specific geography or population. Examples include platforms leveraging GenAI to assist with benefits applications. However, these are often limited in scope (e.g., only focused on paid leave policies) and are not integrated into states' application processes. Other human services program delivery examples include social or assistive robots and GenAI-enabled interview simulations.[580]
3. **Low adoption is driven in part by reliance on pro bono efforts or other non-profit collaborations:** Public sector service delivery agencies leverage external support for pilots that often are for lower risk use cases to reduce administrative burden or to support research. One example is the Illinois Department of Employment Security, which partnered with U.S. Digital Response to improve its translation of unemployment insurance policies using ML-based translation software.[581] Additionally, HHS has seen more

[571] https://nj.gov/humanservices/wfnj/apply/, https://www.mass.gov/info-details/program-verifications-what-information-you-need-to-provide For instance, TANF (or state equivalent program) asks applicants for materials including state ID, social security cards, proof of residency, pay stubs, work hours verification, birth certificates, and marriage certificates.

[572] https://napawash.org/academy-studies/modernizing-public-benefits-delivery-how-innovation-can-deliver-results-for-eligible-households-and-taxpayers

[573] https://www.iiba.org/business-analysis-blogs/how-ai-is-rewriting-the-rules-of-data-analysis/

[574] https://napawash.org/academy-studies/modernizing-public-benefits-delivery-how-innovation-can-deliver-results-for-eligible-households-and-taxpayers

[575] https://initiatives.weforum.org/ai-governance-alliance/home International examples include the World Economic Forum's AI Governance Alliance which has brought together stakeholders from 463 public, private, and social sector entities to share knowledge of AI governance best practices.

[576] https://www.usdigitalresponse.org/services/public-sector-generative-ai

[577] https://www.cpsai.org/

[578] https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition

[579] https://www.acf.hhs.gov/opre/report/options-opportunities-address-mitigate-existing-potential-risks-promote-benefits. Based also on focus groups and conversations ACF has had with human service delivery agencies and industry input on AI integrations in human and health services.

[580] https://pmc.ncbi.nlm.nih.gov/articles/PMC10474924/

[581] https://www.usdigitalresponse.org/services/public-sector-generative-ai

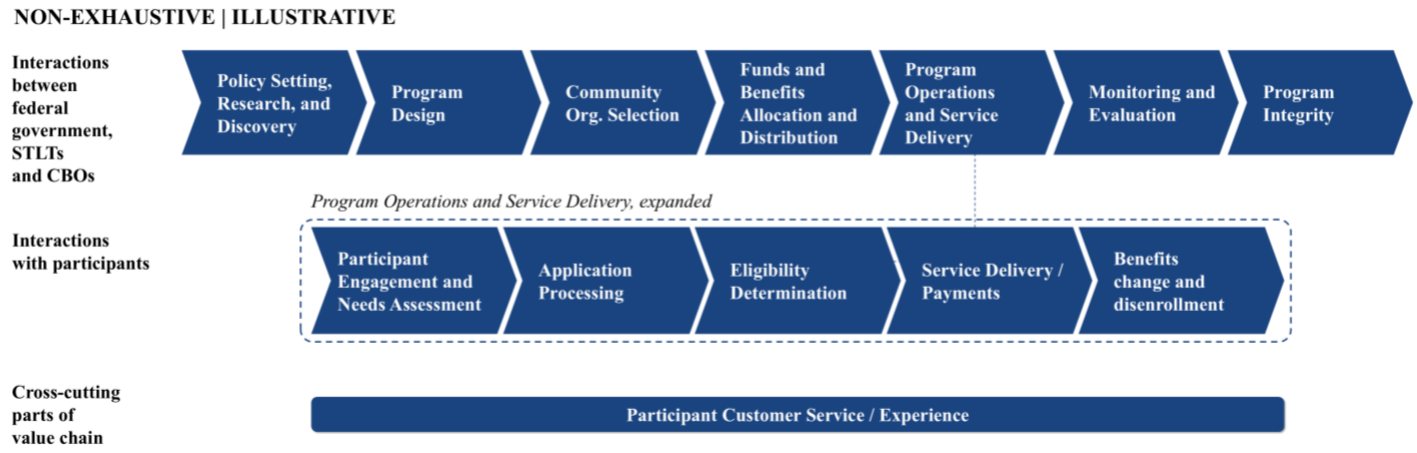
AI-related research activity in human services fields with incentives to develop innovative approaches, such as in child welfare, to prevent the mistreatment of children.[582]

4. **Concerns over potential negative impact of AI may limit adoption in human services:** Stakeholders in human services are reticent to adopt AI tools without risk assessments and stringent requirements to account for potential adverse effects. These are important safeguards as data-driven bias further perpetuates existing inequities, placing served populations at risk of worsened outcomes and further exclusion.[583, 584] Furthermore, as stipulated in the Plan for Responsible Use of AI in Public Benefits, there are rights- and safety-impacting risks from AI applications, such as automated denial of program applications.[585] To account for these risks, stakeholders may place a higher bar on technology vendors and service partners, which, while important, may contribute to a lag in AI adoption in human services compared to other sectors. Trade-offs should be considered

4.5 Potential Use Cases and Risks for AI in Human Services Delivery

Value chains vary across programs and organizations in human services delivery. For example, the type and sequence of activities involved in runaway homeless youth programs, Head Start programs, refugee resettlement, and child welfare services are unique to each program. Below is a general view of the core *functions* underlying human services.

Exhibit 11: Human Services Delivery Value Chain



4.5.1 AI Use Cases Along the Human Services Delivery Value Chain

In the tables below, HHS highlights a non-exhaustive list of potential benefits and risks of AI across the human services delivery value chain. Please note that the use cases detailed below highlight existing or potential ways that AI can be used by a variety of stakeholders in this domain. For details on how HHS and its divisions are using AI, please reference the HHS AI Use Case Inventory 2024.[586]

HHS notes that AI is one technological tool among several for human services delivery stakeholders and that overreliance on AI may pose risks that need to be fully addressed.[587] Further, many technologies (e.g., LLMs) are

[582] https://www.acf.hhs.gov/opre/report/options-opportunities-address-mitigate-existing-potential-risks-promote-benefits
[583] https://pmc.ncbi.nlm.nih.gov/articles/PMC9976641/
[584] https://www.rockefellerfoundation.org/insights/perspective/putting-the-needs-of-vulnerable-populations-first-collaborating-to-address-ai-bias/
[585] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf
[586] https://www.healthit.gov/hhs-ai-usecases
[587] https://www.healthaffairs.org/content/forefront/discrimination-artificial-intelligence-commercial-electronic-health-record-case-study

still being evaluated for potential risks in human services settings. HHS has previously addressed risk considerations related to using AI in other documents, including the HHS Trustworthy AI Playbook and the Plan for Responsible Use of AI in Public Benefits. In addition to the potential use cases, the Department has included potential risks to consider. This list is also non-exhaustive. HHS will consider mitigation steps to address identified risks in the actions proposed later in this chapter (see Action Plan).

Interactions between the federal government, STLTs, and community organizations:

Functional component 1: Policy setting, research, and discovery

The federal government and STLTs establish policies and regulations that guide, inform, and govern all stages of the value chain

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Tools that synthesize multiple, varied datasets to inform policy assessment and creation *E.g., data-driven measurement analytics* AI-driven insight generation from program measurement data, population statistics, and other areas to inform policy setting. This set of tools can synthesize multiple, varied datasets to inform policy assessment and new policy-setting.[588, 589]	**Potential for third-party risk** *E.g., program data breach through a third-party vendor* Third-party data storage could be an access point for a data breach of sensitive population data.[590]

Functional component 2: Program design

The federal government, STLTs, and community organizations design programs and benefit delivery from a systemic to an individual level

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
AI-driven measurement and data analysis tools to inform program design *E.g., policy measurement analytics* Leverage AI-driven insights from the policy-setting stage to inform best practices for designing and delivering programs[591] *E.g., resource and geospatial mapping* AI can support resource mapping, geospatial analysis of population statistics and needs, and predictive modeling to inform program design and organization selection.[592, 593]	**Potential for explainability and accountability risk** *E.g., directing program resources based on a black box algorithm* AI applications could have flawed inputs and lack appropriate safeguards for users to understand decision-making or training data, resulting in mismatched resources for potential participants.[594]

[588] https://www.sciencedirect.com/science/article/pii/S0740624X20300034. History of algorithmic models being deployed to inform government policy

[589] https://www.apec.org/publications/2022/11/artificial-intelligence-in-economic-policymaking International governments are deploying AI to inform policy and measure impact.

[590] See the Cybersecurity and Critical Infrastructure Protection chapter for more information on third-party and data-breach risks for the health and human services ecosystem.

[591] https://www.science.org/doi/10.1126/science.aao4408 The study is related to placement, but the analogy has possible benefits.

[592] https://www.sciencedirect.com/science/article/pii/S0740624X20300034

[593] https://www.apec.org/publications/2022/11/artificial-intelligence-in-economic-policymaking

[594] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

Functional component 3: Community organization selection

The federal government and STLTs select community organizations to execute programs, distribute benefits, and manage and evaluate existing partners

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Evaluation tools for managing community organization networks and identifying new partners *E.g., natural-language notes processing and analytics* Convert narrative format and voice notes taken by caseworkers during network monitoring into digital data that can be evaluated more efficiently and assessed over time[595] *E.g., web-scraping for community organization benchmarking data* AI-driven web search for rapid rate benchmarking, service availability search, and organization prior history to inform network selection, grant approvals, and negotiations[596]	**Potential for misrepresentation due to incorrect interpretation of unstructured data** *E.g., inaccurate structuring of web-scraping data leading to errors in proposal evaluations* AI-driven web-scraping may inaccurately assess data from community organization websites and public references, leading to creation of false or misleading conclusions that affect grant awards or partner evaluation.

Functional component 4: Funds and benefits allocation and distribution

Federal and state governments distribute funds to states and community organizations for programs

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Tools to evaluate grant applications and distribute resources to areas with the highest need, where funding flexibility is allowed by the programs *E.g., proposal synthesis and evaluation* To enable faster, more informed reading of grant applications for discretionary grants[597] *E.g., predictive analytics for funds shortages* AI-driven assessment of where programs and organizations are under/overutilizing funding to improve the allocation of spending across the human services ecosystem[598, 599]	**Potential for inequitable funding allocation based on incorrect output** *E.g., flawed AI-based algorithmic funding distribution leads to resource shortages in STLTs and CBOs* AI-based algorithm distributes funding based on flawed assessment, which could lead to STLTs and CBOs receiving insufficient resources to conduct programs and distribute benefits[600]

Functional component 5: Program operations and service delivery

The activities range from participant engagement and needs assessment to benefits change and enrollment.

Detailed in Interactions with participants

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Detailed in **Interactions with participants** functional components	*Detailed in* **Interactions with participants** functional components

[595] https://pubmed.ncbi.nlm.nih.gov/39396164/. Discussion on use in Healthcare and Public Health

[596] Vendor solutions available for web-scraping.

[597] Commercial tools widely available (e.g., ChatGPT and Bard) with enterprise solutions to build bespoke solutions with private data.

[598] https://www.sciencedirect.com/science/article/pii/S0740624X20300034

[599] https://www.apec.org/publications/2022/11/artificial-intelligence-in-economic-policymaking

[600] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

Functional component 6: Monitoring and evaluation

Assess the effectiveness of individual programs and overall policies in achieving their aims. Identify areas of improvement and recommendations in the future

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Tools to support the evaluation of program effectiveness using unstructured data *E.g., natural-language notes processing and analytics* Convert narrative format notes taken during program delivery (e.g., paper records in homeless shelters, case notes in behavioral health consultations) into digital records for measurement and evaluation[601] *E.g., data-driven performance assessment and reporting* AI-driven data analytics based on data collected from programs, caseworker notes, and external sources to measure program outputs and outcomes, enhancing insights beyond descriptive or leading indicators[602,603]	**Potential for misrepresentation for incorrect output** *E.g., confabulation from AI-generated notes leading to errors in caseworker evaluations* AI confabulation when transcribing caseworker program notes can lead to data misclassification or incorrect measurement of program performance, which can affect policy decisions.[604]

Functional component 7: Program integrity

Ensure accurate, secure, and efficient program delivery and that stakeholders in the value chain are fulfilling their roles. Protect against potential fraud, waste, and abuse

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Data-driven continuous monitoring of the human services portfolio for irregularities *E.g., fraud detection and prevention* Detect irregular patterns in benefits usage, contractor behavior, potential fraud, waste, and abuse using AI-driven analytics versus manual investigation[605] *E.g., automated reporting and insight generation* Data-driven dashboards with the ability to assess program integrity and flag potential irregular activity across a full network of providers[606]	**Potential for incorrect use of AI models and incorrect output** *E.g., improper adaptation of AI fraud detection leading to incorrect program investigation* AI applications developed for one purpose (e.g., fraud detection in financial services, payment processing, or verification) used to serve similar functions in human services (e.g., program fraud detection) and leading to erroneous fraud investigations[607, 608]

[601] https://pubmed.ncbi.nlm.nih.gov/39396164/ Discussion on use in Healthcare and Public Health.
[602] https://www.sciencedirect.com/science/article/pii/S0740624X20300034
[603] https://www.apec.org/publications/2022/11/artificial-intelligence-in-economic-policymaking
[604] https://pubmed.ncbi.nlm.nih.gov/39405325/
[605] https://www.brookings.edu/articles/using-ai-and-machine-learning-to-reduce-government-fraud/
[606] https://learn.microsoft.com/en-us/power-bi/create-reports/sample-artificial-intelligence
[607] https://www.brookings.edu/articles/using-ai-and-machine-learning-to-reduce-government-fraud/
[608] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

Interactions with participants:

Functional component 8: Participant engagement and needs assessment

Create awareness and initiate relationships with potential participants. Assess the needs of individuals or populations and make a preliminary determination of potentially applicable programs

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Assistance tools for the human services workforce to better predict population needs and communicate with participants *E.g., predictive analytics and risk stratification* Predict high-risk individuals and populations and reach out sooner for enrollment, interventions, and wraparound services (e.g., mental health crisis support). Enable caseworkers to flag specific cases for review and personalized treatment[609, 610, 611] *E.g., live-language and cross-cultural translation for caseworkers* AI-driven live translation tools enable caseworkers to interact with participants who speak a different language or caseworkers who speak another language with non-native fluency. Enhancements to translation tools may further assist in identifying cross-cultural communication barriers extending beyond language (e.g., non-verbal communication and cultural practices)[612]	**Potential for incorrect output** *E.g., inaccurate live translation* AI-powered live translation incorrectly communicates information between participants and caseworkers, leading to critical gaps in communication, especially when discussing legal documents or care[613, 614]

[609] https://www.medicaid.gov/state-resource-center/innovation-accelerator-program/iap-downloads/program-areas/factsheet-riskstratification.pdf

[610] https://www.ajmc.com/view/improving-risk-stratification-using-ai-and-social-determinants-of-health

[611] https://www.ncbi.nlm.nih.gov/books/NBK475995/

[612] Multiple vendors exist alongside publicly available solutions like Google Translate.

[613] https://www.sciencedirect.com/science/article/pii/S2772941924000012 Live translation has been shown to outperform machine translation, but performance is still not fully accurate or adaptable to nuanced cultural differences.

[614] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

Functional component 9: Application processing

Collect required data and documentation from other agencies (where possible), potential clients and participants, or their caregivers, and process benefits applications

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Platforms to process applications more rapidly and accurately and provide plain-language information to participants *E.g., predictive eligibility determination* Enable people to understand what programs are available to them with a strong indication of eligibility based on a limited set of demographic and social factors. Further, partially complete the application process based on simplified data and articulate how to finish the process with plain language[615] *E.g., streamlined application processing* Automate application tasks where possible and use data connections from multiple sources and agencies to auto-fill applications and accelerate decision-making. Further, enables interoperability to process multiple programs with similar or the same streamlined program[616]	**Potential for algorithmic bias in decision-making** *E.g., generating misrepresentative benefit determinations for similarly situated people* If AI is applied to aspects of human services delivery, including program eligibility determination, fraud detection, or risk-stratification, there is a risk that those programs will misclassify populations and individuals based on historical misrepresentation in underlying data, and influence decisions in prejudicial ways.[617, 618]

Functional component 10: Eligibility determination

Determine whether a person is eligible for the program or benefits they have applied for and for what level of support

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Connect across multiple disparate human services systems to improve benefit selection and speed up service delivery *E.g., outcomes and follow-on services prediction* Predict the likelihood that individuals enrolling in one program will likely be eligible for and could use another (e.g., X% of enrollees in SNAP are likely to require other cash assistance) and recommend those services using plain language at the time of enrollment[619, 620] *E.g., integrated benefits delivery systems* AI-driven integration of data, applications, eligibility determination, and service delivery across programs in multiple agencies (e.g., across healthcare, human services, housing, and food security) [621]	*See risks outlined in Functional component 8: Application processing*

615 https://www.thomsonreuters.com/en-us/posts/corporates/ai-family-leave-law/

616 https://europepmc.org/article/pmc/pmc10114030 Efforts to streamline or automate the prior authorization process could be applied in other health and human services areas like public benefits.

617 https://jswve.org/volume-20/issue-2/item-05/

618 https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

619 https://pmc.ncbi.nlm.nih.gov/articles/PMC7125114/ Studies done in healthcare settings to predict outcomes or need for follow-on service (e.g., re-admission).

620 https://pmc.ncbi.nlm.nih.gov/articles/PMC11161909/ Recently published survey of studies on the use of AI to predict outcomes.

621 https://pmc.ncbi.nlm.nih.gov/articles/PMC9723913/

Functional component 11: Service delivery and payments

Provide services or payment to participants based on their eligibility. This part of the value chain may include multiple steps and services but is simplified here

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
AI-generated service content and AI-supported platforms to increase the reach and effectiveness of programs *E.g., AI-generated service content* Guidance for and promotion of AI-supported platforms that fulfill the goals of HHS agencies (e.g., social isolation games for the elderly population, digital therapeutic interventions like chatbots for cognitive behavioral therapy, conversational agents for mental health programs) [622] *E.g., AI-enabled robotics in elderly or disability care* Social robots can help treat isolation or dementia in elderly populations. Assistive robots can help with daily tasks, including personal hygiene and mobility.[623]	**Potential for bias or incorrect output** *E.g., improper assessment of program participant suitability for an education program* Biased data or flawed algorithms are used to determine eligibility or conditions for workforce training programs, leading to incorrect placement or program offers.[624] *E.g., misaligned assignment of caseworkers* Flawed AI-driven assessment of case complexity could exacerbate workforce challenges through misallocated resources.[625]

Functional component 12: Benefits change and disenrollment

Renew and update recipient benefits or disenroll participants when they no longer meet assistance criteria

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Proactive enrollee management to ensure accurate re-enrollment and benefit changes *E.g., enrollee address and information verification* Use AI to confirm enrollee information and assist with confirming eligibility, disenrolling, or reenrolling. Tool relevant during determination windows and when the participant may have had a change in life event[626] *E.g., proactive eligibility change notification* Use of data integrated across multiple agencies to predict when participant eligibility (e.g., benefits cliffs) will change and proactively notify using plain language[627]	**Potential to magnify participant trust concerns and AI skepticism** *E.g., overcollection of data* The overcollection of data (or perception of overcollection or misuse) for an AI model predicting benefits change (e.g., loss of benefits) may enhance distrust, particularly for underrepresented populations who may already have negative perceptions of human services programs.[628]

[622] Private mental health technology companies are using AI to generate content for mental health programs.

[623] https://pmc.ncbi.nlm.nih.gov/articles/PMC10474924/

[624] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

[625] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

[626] Multiple vendor solutions across other sectors (e.g., financial services) exist.

[627] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

[628] https://www.ama-assn.org/system/files/ama-patient-data-privacy-survey-results.pdf

Cross-cutting parts of the value chain:

Functional component 13: Customer service/experience

Provide customer support and information to people as they navigate the process from needs assessment to service delivery and benefits change

Potential use cases (non-exhaustive)	Potential risks (non-exhaustive)
Customer support tools to improve interactions of human services staff and more simply and accurately offer support to participants *E.g., enhanced external chatbot or virtual assistant* Create GenAI-enabled chatbots or virtual assistants that can answer questions for participants or potential applicants in plain language in multiple languages. Assist with basic eligibility prediction and integration into application processing[629] *E.g., synthesized participant feedback tool* Use GenAI and connection to unstructured caseworker notes and call center feedback to conduct sentiment analysis and identify trends and common themes from participant inquiries and calls to human services call centers[630] *E.g., community organization and STLT-facing chatbot or virtual assistant* GenAI-enabled chatbot or virtual assistant for community organizations and STLTs to understand grant and award requirements, answer questions related to new policies, and receive direction related to programmatic questions[631] *E.g., back-end call center optimization* AI-driven analytics to understand what service channels, times, and other circumstances require differing capacity and optimizing workforce to accommodate demand[632]	**Potential for incorrect output** *E.g., internal staff chatbot or external facing chatbot providing false information* Internal AI-powered chatbots used by support staff to interact with participants could provide human services staff with incorrect information to provide to participants, potentially reducing benefits access.[633] A similar public-facing chatbot could similarly create risks if it provides incorrect information, creates barriers to program access, or leads participants to believe they are ineligible for programs.

4.6 Action Plan

In light of the evolving AI landscape in human services delivery, HHS has taken multiple steps across issuing new guidelines for STLT use of AI in public benefits, practice sharing through public-private partnerships, and provision of grant funding to promote responsible AI. The Action Plan below follows the four goals that support HHS's AI strategy: 1. catalyzing health AI innovation and adoption; 2. promoting trustworthy AI development and ethical and responsible use; 3. democratizing AI technologies and resources; and 4. cultivating AI-empowered workforces and organization cultures. For each goal, the Action Plan provides context, an overview of HHS and relevant other federal actions to date, and specific near- and long-term priorities HHS will take. HHS recognizes that this Action Plan will require revisions over time as technologies evolve and is committed to providing structure and flexibility to ensure longstanding impact.

[629] https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2023.1275127/full

[630] Companies developing GPTs and LLMs offer enterprise solutions to tailor their GenAI tool to specific organizational needs.

[631] https://journals.sagepub.com/doi/10.1177/02750740231200522

[632] https://www.nber.org/papers/w31161 The paper assesses the impact of AI in customer support roles in general.

[633] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

4.6.1 Catalyze AI Innovation and Adoption

HHS could promote AI innovation and adoption through opportunities related to the following areas:

1. Unlocking resources for AI adoption and modernizing IT and tech infrastructure
2. Ensuring data quality and availability for AI adoption

Below, the Department discusses context, HHS actions to date, HHS near-term priorities, and potential long-term actions.

1. Unlocking resources for AI adoption and modernizing IT and tech infrastructure

Context:

Grants and contracts in human services do not tend to allocate funds for AI-related investments, nor do they require demonstrated IT capabilities as conditions for awards. Overall, the sector faces funding and workforce shortages that leave many stakeholders feeling unable to meet demand for their services over a year.[634] STLTs and community organizations may lack funding that can be directed toward investments in AI or improving tech infrastructure.[635] As a result of this persistent funding shortage, non-profits spend less on IT infrastructure than the private sector, even though more technologically advanced non-profits are more likely to fulfill their missions.[636] This lack of investment has led to outdated IT infrastructure in agencies and community organizations or overreliance on analog and paper record keeping. Organizations may require leapfrogging several IT maturation stages to incorporate AI into their operations. STLTs and CBOs seeking to make transformational investments in IT without a proper technological foundation may face additional challenges, including reduced service quality due to the need to troubleshoot AI use. Greater resources for AI adoption would enable multiple opportunities, including enhancing interoperability, increasing the pace and quality of funds distribution as well as improving service quality and experience.

[634] https://nff.org/learn/survey A 2022 survey of non-profits from the Nonprofit Finance Fund found that more than half of participating organizations felt they would be unable to meet the demand for their services in the upcoming year.

[635] https://pmc.ncbi.nlm.nih.gov/articles/PMC6816239/

[636] https://ssir.org/articles/entry/taking_on_tech_governance#

HHS actions to date (non-exhaustive):

- **HHS's Plan for the Responsible Use of AI in Public Benefits:**
  - Outlined additional areas of support for STLTs about promoting AI use in public benefits, including providing information on funding available to STLTs.
  - Recommended specific enablers for the effective adoption of AI in public benefits administration among STLTs and vendors. These enablers include improved IT infrastructure, high-quality data, and appropriate safeguards.
  - Explored providing technical assistance to STLTs attempting to implement responsible AI in public benefits to increase their capacity to utilize AI appropriately and root out and mitigate risks.
- **Leveraged existing partnerships and developed new relationships to coordinate the promotion and adoption of AI.** HHS has existing partnerships with coalitions, advisory committees, and other organizations and is creating new relationships to help share best practices and lessons, including mistakes, across jurisdictions. This information sharing can shorten the learning curve for newer adopters and provide hands-on, tactical guidelines, including templates for policies, governance, and the procurement of AI tools.
- **Provided grant funding to CBOs, improving service quality through AI applications.** ACF and ACL are funding organizations using AI to improve their operations or program delivery in multiple ways, including using robotics in assisted living facilities or launching AI-enabled chatbots.[637, 638]

Other federal actions to date (non-exhaustive)

- **USDA released the Framework for STLT Use of Artificial Intelligence in Public Benefits in April 2024 (referred to as the "USDA's Framework for STLT Use of AI").**[639] Mirroring HHS's Plan for Responsible Use of AI in Public Benefits, the Department of Agriculture's plan provides guidelines for STLT's use of AI, including determining the benefits and goals of AI adoption, interactions with vendors, and the responsible use of data and IT system design.

HHS near-term priorities:

- Identify funding opportunities available to lower-resourced STLTs and community organizations for AI adoption in human services, including IT modernization, data quality improvement, or other investment-type grant programs.
- Explore private sector collaborations that could provide technical assistance to HHS, STLTs, and community organizations interested in adopting AI applications and modernizing IT for their human services programs.
- Compile and make available best practice implementations of IT modernization, including for those categories outlined in the Plan for Responsible Use of AI in Public Benefits (e.g., IT readiness, and best practices interoperability) in the human services delivery ecosystem.
- Explore expanding the procurement guide for STLTs (above what was provided in the **Plan for Responsible Use of AI in Public Benefits**) to use as they evaluate AI tools in their information systems that help administer public benefit programs.

[637] https://acl.gov/news-and-events/announcements/acl-awards-20-field-initiated-projects-program-grants

[638] https://acl.gov/news-and-events/announcements/new-funding-opportunity-small-business-innovation-research-program-4

[639] https://www.fns.usda.gov/framework-artificial-intelligence-public-benefit

HHS long-term priorities:

- Explore resources for government, non-profit, and research collaborations working in the human services ecosystem to adopt AI for improving their programs and benefits.
- Identify best-practice open-source AI and infrastructure tools for human services organizations to leverage mapping tools to specific high-value use cases.
- Evaluate opportunities to modernize HHS's IT infrastructure to support greater AI adoption in the human services ecosystem.
- Integrate resource and technical assistance opportunities into mechanisms such as block grants, advanced planning documents, challenge grants, and federal contracts for AI applications that address human services programs dependent on resourcing and where most appropriate and feasible (e.g., promote health and well-being).
- Consider designing "moonshot" competitions such as those used by CMMI, GSA, and Defense Advanced Research Projects Agency (DARPA) for system-level human service delivery solutions and providing resources and assistance for promising solutions to scale.

2. Ensuring data quality and availability for AI adoption

Context:

Owing in part to legacy IT, program requirements, concerns over participant privacy, and employee/client digital literacy (among other factors), many human services agencies record data in unstructured, non-standardized formats. These data are difficult to incorporate into AI-driven applications. With improved data quality, governance, and interoperability, HHS could drive greater adoption of AI use cases that require accessible data. Additionally, AI itself can improve data availability through better interpretation of unstructured information. Data quality and availability improvements may enhance interoperability for service coordination and move the Department closer to its goal of a human-centered approach that seamlessly connects participants' platforms to programs across human services, healthcare, and other public benefits.

HHS actions to date (non-exhaustive):

- **The Plan for Responsible Use of AI in Public Benefits (HHS)** recommended enablers for the effective adoption of AI among STLTs, including improving data quality and access.

HHS near-term priorities:

- Continue to issue guidelines and establish interoperability standards where authorized for sharing data across programs, departments, levels of government, and community organizations.
- Identify, with STLT and community organization input, priority areas of human services delivery with gaps in data quality and collection (e.g., translations for less widely spoken languages) and align on a path forward for improvement.
- Promote data quality standards, governance, and access to best practices observed in the human services ecosystem or adjacent areas with adaptations to human services, including best-practice for AI use to improve data-processing and structuring.
- Explore private sector collaborations that could provide technical assistance to HHS, STLTs, and community organizations interested in improving data quality.

HHS long-term priorities:

- Consider implementing shared sandbox environments to accelerate piloting use cases and reduce the cost of understanding the return on investment in AI applications.
- Review HHS-owned datasets for quality and applicability to AI use cases and create improvement plans where necessary.

4.6.2 Promote Trustworthy AI Development and Ethical and Responsible Use

HHS could ensure that AI use remains trustworthy and safe by:

1. Providing guidelines on balancing risks to served populations and establishing participant trust with opportunities for AI applications.

Below, the Department discusses the context, HHS actions to date, HHS near-term priorities, and potential long-term actions.

1. Providing guidance to served populations on balancing risks with opportunities for AI applications and establishing participant trust

Context:

Many stakeholders are vocal and active in ensuring that the populations that HHS serves are directly involved in developing AI applications and determining data used in AI models.[640] Tailoring AI in human services to match the needs and cultural context of participants could improve service quality and accessibility of services. This is especially important for populations that have historically been under- or misrepresented in data (e.g., refugees, tribal communities, and people with disabilities) and for AI applications that could affect peoples' rights and safety (e.g., benefit eligibility determination).[641] Initial research on AI indicates that cultural context and background impact preferences for using AI.[642] However, more time and investment would be required to fully allay concerns about misrepresenting served groups in AI applications.

Further, concerns for participant data privacy and safety may inhibit technology adoption. For instance, among AI vendors, there is broad awareness of federal AI risk frameworks and support for guardrails; however, many offerings do not provide the level of transparency required to assuage human service stakeholder concerns. Human services agencies attempting to address data bias and privacy concerns without sufficient guidelines may disqualify AI solution vendors unwilling to offer additional transparency on their training data.

Finally, HHS's existing authority enforcing the use of AI in human services delivery is limited. It lacks a robust approach to AI oversight, including certifications, privacy and security controls, and third-party evaluations. Clear risk assessment and mitigation standards for organizations that develop and deploy AI in human services settings could mitigate the risk of inappropriate use.

[640] https://datasociety.net/library/democratizing-ai-principles-for-meaningful-public-participation/, https://www.upwardlyglobal.org/ai-for-impact-report/

[641] https://www.ncbi.nlm.nih.gov/books/NBK584407/

[642] https://hai.stanford.edu/news/how-culture-shapes-what-people-want-ai

HHS actions to date (non-exhaustive):

- **ACF Policy on Generative AI Tools from July 2024**[643] provided principles to encourage the appropriate and responsible use of GenAI to support its workforce and improve service delivery. These requirements for ACF staff and contractors include understanding the tool's purpose and limitations, understanding how to securely use and protect participant data, reviewing and fact-checking the output, and being transparent with use.
- **Plan for Responsible Use of AI in Public Benefits (HHS):**
    - Provided recommendations for Managing Risks for the Use of Automated Algorithmic Systems, including those focused on managing the highest risk AI use cases, ensuring safety and security, sustaining human judgment, allowing participant opt-outs, protecting recipient interests, and safeguarding civil liberties.
    - Recommended establishing effective governance mechanisms for AI risks consistent with the six principles outlined in the NIST AI Risk Management Framework. Additional recommendations include maintaining an inventory of automated and algorithm-based technologies, creating a formalized process to evaluate risks in AI use, and educating vendors about their AI governance practices.
- **HHS Trustworthy AI Playbook (2021) outlined guidelines for the internal use of AI applications at HHS.**[644] It provides guidelines to ensure that AI applications internal to HHS are developed and deployed ethically, effectively, and securely, aligned with federal standards, and promote public trust throughout the AI life cycle. However, these guidelines have not been tailored specifically to human services programs and are mostly limited to guidelines on internal use cases rather than promoting external adoption.
- **Joint Statement on Enforcement of Civil Rights, Fair Competition, Consumer Protection, and Equal Opportunity Laws in Automated Systems (April 2024)**[645] clarified the ability to use enforcement action for violations from automated systems and that it can use that authority to enforce rules related to equity, including enforcing Civil Rights, Fair Competition, and Consumer Protection.
- **Published internal governance documents for external reference.** ACF has published its AI Activation Toolkit.[646]

Other federal actions (non-exhaustive):

- **USDA's Framework for STLT Use of AI** includes most risk management and governance recommendations from the HHS Plan for Responsible Use of AI in Public Benefits. This reflects the overlapping responsibility of many STLT human services departments to administer a mix of HHS and USDA programs (e.g., TANF and SNAP).

HHS near-term priorities:

- Issue new guidelines and recommendations as outlined in the **Plan for Responsible Use of AI in Public Benefits,** including clarifying principles for roles of human intervention in automated systems, customer support, and GenAI use.
- Consider issuing guidelines on best-practice interactions with participants to explain and establish trust in using AI in human services programs.
- Research effective methods for using AI in human services while adopting best-practice safety standards (e.g., bias mitigation and maintaining human-in-the-loop).
- Define applicable regulatory authorities for using AI in human services delivery (e.g., for AI-enabled devices in assisted and community living) and clarify HHS's role in enforcement regarding the trustworthiness and safety of AI use.

[643] https://www.acf.hhs.gov/sites/default/files/documents/main/ACF-Generative-AI-Policy-June-2024.pdf

[644] https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf

[645] https://www.justice.gov/crt/media/1346821/dl?inline

[646] https://www.acf.hhs.gov/ai-data-research/artificial-intelligence-acf

HHS long-term priorities:

- Integrate AI safety and transparency requirements into HHS funding mechanisms by complying with best-practice guidelines for block grant conditions, state plans, advanced planning documents, challenge grants, and federal contracts in coordination with relevant federal partners.
- Explore direct resource support opportunities for HHS, STLTs, and community organizations to monitor AI applications and use risks.

4.6.3 **Democratize access to AI technologies and resources across the U.S.,** including for underrepresented populations

HHS could ensure equitable access to AI through actions related to several opportunities, including:

1. Raising the floor of constituent digital literacy and digital penetration
2. Identifying areas of cooperation across sectors to improve AI-related economies of scale

Below, the Department discusses the context, HHS actions to date, HHS near-term priorities, and potential long-term actions.

1. Raising the floor of constituent digital literacy and digital penetration

Context:

In some programs, the population likely to access human services programs is older and less likely to speak English proficiently. Historically, these populations have lower digital literacy, internet access, and smartphone penetration rates.[647] Further, 24 million Americans, some of whom overlap with human services populations, lack access to broadband internet.[648] This "digital divide" limits the effectiveness and solution space for client-facing AI applications in human services. AI applications could mitigate these challenges; however, it requires a baseline digital literacy and capability that some parts of the human services ecosystem may not have.

One additional consequence of the digital divide concerns data access and consent. An agency or community organization may be unable to obtain data for populations with limited digital access or who cannot or will not consent to sharing their data. Data gaps can exacerbate data quality issues and hinder the deployment of equitable and contextualized predictive analytics and the development of better AI tools. Increasing the connectivity and digital experience of potential participants could increase their access to services that otherwise required technology participants did not previously have or understand.

HHS actions to date (non-exhaustive):

- **The Plan for the Responsible Use of AI in Public Benefits** engaged with the public to collect broad feedback on the use of AI in public benefits. These engagements included listening sessions, advisory committees, tribal consultations, webinars, workshops, and other activities intended to include more voices in HHS AI-related policy development.

Other federal actions (non-exhaustive):

- **The USDA Framework for STLT Use of AI** closely mirrors recommendations from HHS's plan for ensuring equitable access and protecting against bias in using AI in public systems.

[647] https://www.ntia.gov/blog/2022/switched-why-are-one-five-us-households-not-online

[648] https://www.ntia.gov/blog/2022/switched-why-are-one-five-us-households-not-online

HHS near-term priorities:

- Establish ongoing consultation channels with the inclusion of various partners, such as IT/AI collaboratives, community-based groups, AI subject matter experts, research organizations, frontline staff, and participants and representatives across different populations (e.g., urban/rural, children/adults, older adults, people with disabilities) and backgrounds (e.g., race, sexual orientation, ethnicities, and language) to identify paths to improve AI accessibility in human services.
- Develop guidelines for how STLTs and community organizations can address inequities in digital literacy in populations they serve, including guidelines for identifying the historical, contemporary, and structural contributors to the inequities that drive disparities in AI adoption.
- Share information on HHS-implemented AI use cases to model opportunities for human services organizations.
- Compile and research potential AI use cases to mitigate or address inequities.

HHS long-term priorities:

- Consider establishing grant or assistance programs where authorized and resourced to address inequities in access to impactful AI in the human services ecosystem (e.g., awards for populations with a high digital divide).
- In coordination with appropriate entities, explore developing and implementing education campaigns for at-risk demographic groups focused on the harms of AI-enabled scams (e.g., deepfake-supported scams like impersonation of family members, government officials, and financial institutions) that are intended to defraud individuals of money and resources.
- Continue to evaluate and support methods to ensure underserved populations may access and benefit from AI.

2. Identifying areas of cooperation across sectors to improve AI-related economies of scale

Context:

Even where agencies and organizations find they have funds to invest in AI applications for their own organization, they may face two additional accessibility barriers. First, a smaller organization may find it challenging to capture the benefits of AI at scale without broader sector wide investment beyond its means.[649] For instance, a use-case like fraud detection or program measurement analytics may require data or technical capabilities from across multiple organizations. Second, a small organization may lack an in-house workforce with enough technical expertise to evaluate vendor options and integrate new solutions (for further details on opportunities related to the workforce, please see the next section on "Cultivating AI-Empowered Workforces and Organizational Cultures").[650] Thus, the size and makeup of these organizations may hinder access even where investment exists. Addressing challenges with scale would improve service experience and quality through greater access to program-enhancing AI. Likewise, it could increase accessibility of services where under-resourced STLTs and community organizations are able to reach more people through AI-enabled platforms that they otherwise lacked the scale to adopt.

HHS near-term priorities:

- Identify use cases or IT investments that are the most promising for the human services delivery ecosystem but require scale beyond community organizations or STLTs (e.g., fraud detection capabilities).
- Consider grant and technical assistance opportunities for deploying use cases requiring coordinated activity or larger scale.

[649] https://hbr.org/2022/03/how-to-scale-ai-in-your-organization

[650] https://www.salesforce.com/news/stories/public-sector-ai-statistics/

HHS long-term priorities:

- Explore creating an "AI for human services" toolkit with critical resources on AI adoption in human services and making it open source to STLTs and community action organizations.
- Consider convening an HHS AI center of excellence team that provides technical expertise and capabilities to HHS, STLTs, and community organizations and develops their capabilities for the Plan's goals.

4.6.4 Cultivating AI-Empowered Workforces and Organizational Cultures

HHS could cultivate AI-empowered workforces and organizational cultures through actions related to several opportunities, including:

1. Improving human services employee digital literacy, talent, and openness to adopting technology
2. Using AI to mitigate the labor workforce shortage in human services

Below, the Department discusses the context, HHS actions to date, HHS near-term priorities, and potential long-term actions.

1. Improving human services employee's digital literacy, talent, and openness to adopting technologies

Context:

Through informal conversations with human services stakeholders, HHS has heard both concerns about the effects of AI on their workforce and requests for assistance in educating the workforce on AI. These concerns correspond to an overall shortage of AI expertise in the public sector that could impede adoption.[651] Additionally, in response to an HHS RFI on AI in human and health services delivery, AI developers and implementers frequently cited organizational readiness (human capacity, technical infrastructure, data quality, change management) as a critical barrier to the successful use of AI.[652] Additionally, vendors require guidelines from AI-informed experts in human services on mission-driven use case identification and prioritization; however, these experts are scarce in many agencies and community organizations. Further, limited in-house digital talent among stakeholders impedes the adoption and enthusiasm for new tools. This gap extends to the most senior roles in non-profits engaged in human service delivery, where boards often lack a member with deep tech experience.[653] Finally, even where agencies or CBOs potentially make large-scale investments in IT, a lack of training on AI tools could leave staff unprepared to use new technologies effectively and potentially reduce service quality. Where human services stakeholders are open to adopting new technologies, they can use AI to enhance the capabilities of their workforce, potentially freeing capacity to serve the growing population who access human services programs.

HHS actions to date (non-exhaustive):

- **The Plan for the Responsible Use of AI in Public Benefits** recommended actions for STLTs to support the workforce in responsibly using AI, including training them on developing and using automated and algorithmic systems, sustaining staff judgment when using AI, and exercising control over algorithmic systems when engaging third-party vendors. USDA's Framework for STLT Use of AI closely mirrors these recommendations.
- **HHS AI Trustworthy AI Playbook (2021)** provided education on AI concerning internal HHS systems. Information included benefits, drawbacks, and potential applications of AI in the Department. The

[651] https://www.salesforce.com/news/stories/public-sector-ai-statistics/

[652] Informal conversations between HHS working group and vendors.

[653] https://ssir.org/articles/entry/taking_on_tech_governance#

Playbook also provides guidelines on incorporating trustworthy AI principles into work routines and overseeing AI-related projects.

HHS near-term priorities:

- Explore direct grant or technical assistance opportunities for workforce training and technical assistance within HHS, among STLTs, and in community organizations.
- Develop best practice guidelines for how federal and state agencies and community organizations can improve AI readiness of their workforces.
- Establish digital literacy and AI literacy training for HHS staff working in human services.
- Support or initiate partnerships between the human services ecosystem and private sector leaders in AI and digital transformation to facilitate information sharing.
- To the extent desired by tribal nations and where resources are available, support tribal nations working to regulate and implement oversight of AI in human services.

HHS long-term priorities:

- Make HHS-internal digital and AI literacy training publicly available for STLTs and community organizations.
- Convene regular AI in human services conferences with learning tracks, practical workshops, and recorded resources.

2. Using AI to mitigate the labor workforce shortage in human services

Context:

As previously noted, the Bureau of Labor Statistics projects a 67,300-person social worker shortage across the U.S. annually over the next decade.[654] A more digital, AI-enabled workforce could identify and deploy use cases that enhance the capabilities of the human services workforce and focus staff on value-added activities and on participant interactions. This could alleviate elements of the job driving low satisfaction and high turnover. However, there are concerns about using AI to augment the human services workforce. First, there are concerns that AI adoption may result in workforce displacement[655] without improving productivity or service quality.[656] Second, overreliance on AI to increase workforce capacity may remove the human element from human services programs for participants. HHS is considering ways to balance these concerns alongside the opportunity for AI in the human services workforce.

Other federal activities (non-exhaustive):

- **The Department of Labor released comprehensive AI Practices (October 2024)** that provide strategies for how AI can benefit workers and businesses while focusing on workers' rights, job quality, well-being, privacy, and economic security.

HHS near-term priorities:

- Share best practices from the human services delivery ecosystem for expanding the workforce's AI capacity.
- Explore additional areas to issue guidelines specific to human services for responsibly adopting AI aligned with **Department of Labor AI Practices** and participants' desires to maintain human interaction.

---

[654] https://www.bls.gov/ooh/community-and-social-service/social-workers.htm

[655] https://www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent.html

[656] https://www.healthaffairs.org/content/forefront/discrimination-artificial-intelligence-commercial-electronic-health-record-case-study AI tool for predicting no-shows can have adverse effect of reducing service quality if biased algorithms incorrectly predict no-show probability, increasing chances that set of individuals are double-booked for appointments.

HHS long-term priorities:

- Review existing guidelines for program delivery and interoperability provided to STLTs to identify areas where AI can alleviate workforce capacity constraints.

4.7 Conclusion

AI has the potential to address underlying challenges in the human services ecosystem, from persistent workforce shortages to low participant satisfaction with programs. Eventually, AI applications may improve human services programs for those who participate in them. However, fundamental challenges have impeded adoption, including a lack of funding and concerns over rights and safety. Despite the challenges, HHS believes that AI can improve program quality, increase access, reduce administrative burden, and enhance interoperability of public benefits systems.

Further, HHS is well positioned to help the human services ecosystem overcome its challenges and realize the benefits of AI. At the same time, it can establish standards and educate the public on the risks inherent in AI, ensuring that AI applications are trustworthy and safe. It can also function as a convener to elevate best practices from the broader ecosystem and highlight lived experiences with AI and its effects on served populations and historically misrepresented groups. This Plan will evolve as the AI landscape changes, but HHS believes that the actions outlined in this Plan will materially advance HHS and the U.S.'s strategic interest in AI.

5 Public Health

5.1 Introduction and Context

For this Strategic Plan, public health is defined as "the science and art of preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society, organizations, public and private communities, and individuals." [657] U.S. public health covers a diverse range of issues like infectious diseases, substance use disorders, non-communicable diseases, environmental health and climate adaptation, and mental and behavioral health. Public health challenges and their underlying disease processes are complex and often involve interactions across biological, social, economic, and other dynamics, requiring collaboration across high-level actors, both public and private.

The COVID-19 pandemic and its aftereffects highlighted severe challenges and gaps in the U.S. public health ecosystem, including (1) difficulty rapidly collecting, sharing, and analyzing information, (2) rising health inequity and public distrust of science, and (3) longstanding resourcing and staffing strains.[658]

AI can help find new solutions to these challenges. For instance, AI can help by automating processes across the data life cycle (e.g., data cleaning, validation, and aggregation) and analyzing vast amounts of data to identify patterns and generate insights, thereby improving public health decisions, interventions, and programs and ensuring resources are allocated where they are needed most. The integration of AI into public health has the potential to significantly enhance disease monitoring and intervention design (e.g., through automated outbreak detection and rapid analysis of large datasets and non-traditional data sources for non-communicable diseases). Additionally, AI can help improve diagnostic accuracy, better engage diverse populations, and optimize the use of public health's often limited resources. Strategic focus and resourcing from HHS agencies, including CDC, NIH, and ASPR, will be critical to driving this transformational change in the public sector. HHS has a unique challenge and opportunity to drive innovation in public health, and by extension private sector healthcare.

Federal activities related to data modernization and AI adoption efforts have been ongoing for nearly a decade across the public health ecosystem. Examples include E.O. 13994, E.O. 13960, and ASTP rules HTI-1 and HTI-2. Activities also include initiatives to increase the availability and quality of data, agency-specific implementation efforts, and federal rules and policies related to a data-driven response to COVID-19 and response readiness for future events. These efforts, while not intended solely for the quick uptake of AI, directly connect and support public health agencies' efforts to be able to deploy AI. The U.S. public health ecosystem is only as strong as its weakest link; without data modernization and interoperability, isolated health entities will not have the means to contribute to and benefit from shared data and AI use. This prevents the entire ecosystem from building the comprehensive data view necessary to effectively detect, understand, and address public health issues. The foundation is being laid to break down silos and encourage the use of AI, and there is an opportunity for CDC and the rest of HHS to be a lighthouse for others in the ecosystem.

AI innovation and usage have also been discussed in almost every global health forum over the last few years and there is opportunity for AI to improve health globally. Multiple HHS actions related to global health have already been launched (e.g., the ARPA-H program on AI antibiotics to combat anti-microbial resistance).[659] However, as HHS representatives stated at the G7 conference, the effectiveness of AI is determined in large part by the strength

[657] https://www.cdc.gov/training-publichealth101/media/pdfs/introduction-to-public-health.pdf
[658] https://www.cdc.gov/workforce/php/about/index.html
[659] https://arpa-h.gov/news-and-events/arpa-h-award-aims-combat-antimicrobial-resistance

of a country's enabling environment—one that is trustworthy, accessible, and free of bias. Countries around the world, including the U.S., are looking at how best to use AI to improve healthcare systems and protect against major health threats wherever they arise.

In addition to its many opportunities, AI use is accompanied by serious risks related to privacy, ethics, and equity, many of which can be further addressed by HHS actions. To maximize the benefits of AI in public health, existing efforts will have to be accelerated and integrated into a cohesive strategy that balances innovation with safety and security. To that end, later in this document, HHS has outlined a set of strategic priorities to catalyze health AI innovation and adoption, ensure AI use is trustworthy and safe, democratize access to AI technologies and knowledge, and support the cultivation of AI-empowered workforces and organizational cultures.

5.1.1 Action Plan Summary

Later in this chapter, HHS articulates proposed actions to advance its four goals for the responsible use of AI in the sector. Below is a summary of the themes of actions within each goal. For full details of proposed actions please see section 5.6 Action Plan.

Key goals that actions support	**Themes of proposed actions *(not exhaustive, see 5.6 Action Plan for more details)***
1. Catalyzing health AI innovation and adoption	• Encouraging research, development of guidelines, and identification of resources to support evidence generation and scale of AI in public health • Modernizing infrastructure necessary to implement AI and support adoption
2. Promoting trustworthy AI development and ethical and responsible use	• Establishing guardrails to help ensure data quality and accuracy • Standardizing data security policies across the public health ecosystem • Advancing AI tools and techniques that consider and assess health equity from end to end
3. Democratizing AI technologies and resources	• Creating an environment that enables data sharing across the public health ecosystem • Supporting AI adoption, development, and collaboration, especially for STLTs and community organizations who may have limited resources • Developing user-friendly, customizable, and open-source AI tools to broaden access and accommodate a diversity of users
4. Cultivating AI-empowered workforces and organization cultures	• Augmenting and supporting the public health workforce to address burnout and attrition • Promoting AI education and community-based AI approaches tailored to each community's unique need

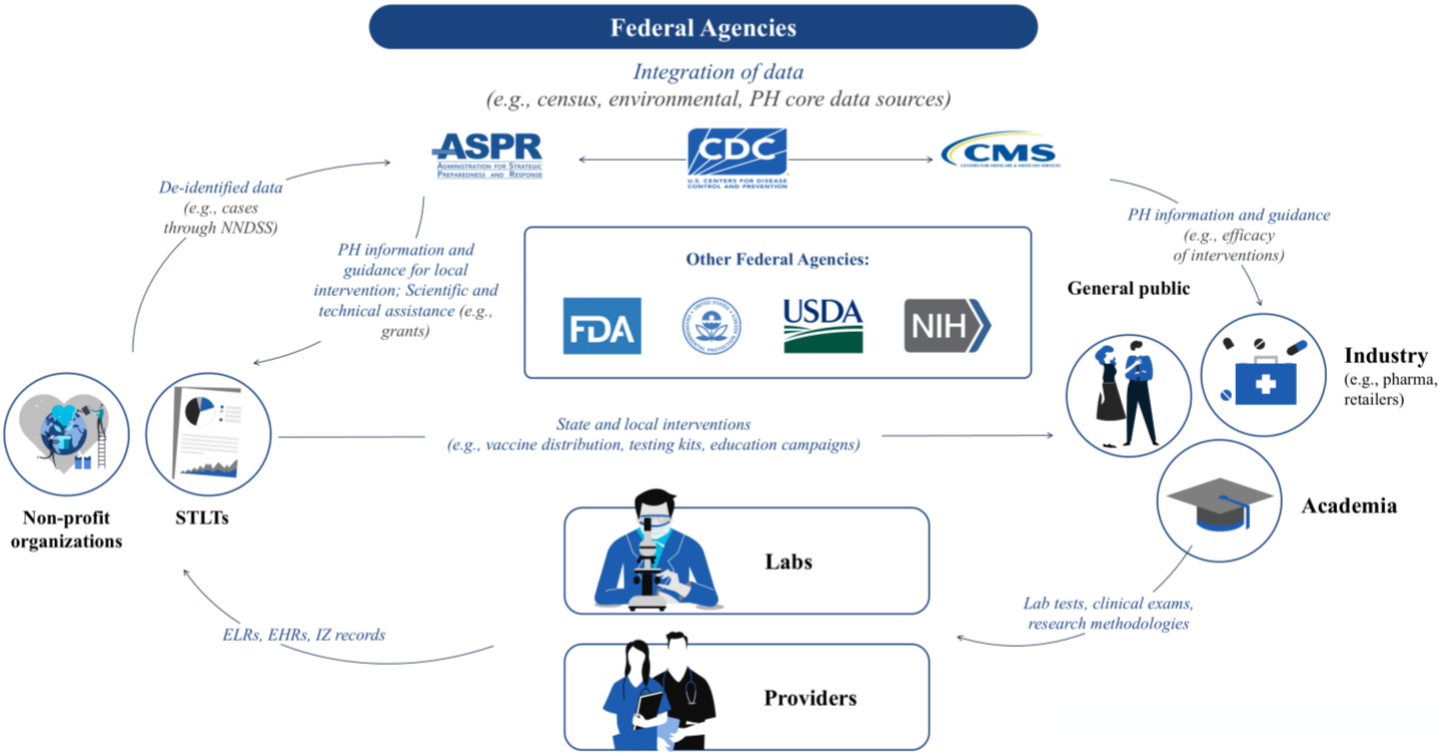
5.2 Stakeholders Engaged in the Public Health AI Value Chain

The U.S. public health ecosystem is anchored on the coordination and support of the federal government and STLTs and relies on the collaboration of a wide range of stakeholders, from providers, health systems, private partners, and researchers to non-profits and the general public to enact positive societal change. To illustrate the diversity of public health actors, below is a non-exhaustive, illustrative diagram of example flows between stakeholders (Exhibit 12) and a bulleted list of stakeholders involved.[660] Please note that neither the diagram nor the list captures all stakeholder roles and interactions. Please refer to other HHS documents for additional details on regulatory guidance and authorities.

[660] Descriptions are illustrative and do not capture the full range of each entity's roles and responsibilities

Exhibit 12: The U.S. Public Health Ecosystem[661]

NON-EXHAUSTIVE | ILLUSTRATIVE



- **HHS agencies:** Public health is supported through the efforts of the operating divisions of HHS, such as:
  - **ACF:** Provides benefits and services to support the well-being of families and children, many of which are related to public health (e.g., behavioral health and abuse prevention).
  - **ACL:** Supports programs for populations with complex needs, particularly older adults and people with disabilities, including nutrition services, medical care, and elder support services.
  - **AHRQ:** Focuses on improving the quality, safety, efficiency, and effectiveness of healthcare for all Americans through research.
  - **ASPR:** Leads national preparedness, response, and recovery from disasters and public health emergencies.
  - **Agency for Toxic Substances and Disease Registry (ATSDR):** Prevents exposure to hazardous substances (e.g., chemicals, pesticides, heavy metals) and mitigates associated health risks. ATSDR conducts risk assessments and health consultations and supports health education.
  - **CDC:** Actively detects, surveils, defines, prevents, and responds to disease outbreaks, administers national health programs, and supports policymaking by providing technical assistance and information.
  - **CMS:** Administers major public healthcare payer programs (e.g., Medicare and Medicaid), outlines conditions of participation related to these programs for healthcare providers contingent on sharing critical public health data (e.g., related to healthcare-associated infections), and could provide payment for specific devices or services.
  - **FDA:** Acts as a core regulator to ensure the safety and effectiveness of medical products and the security of medical devices, including AI-enabled medical devices.
  - **IHS:** Provides a comprehensive healthcare delivery system and ensures culturally appropriate public health and human services are available for American Indian and Alaska Native people to raise the physical, mental, social, and spiritual health of the population to the highest level.

[661] https://www.cdc.gov/public-health-data-strategy/php/about/public-health-ecosystem-data-goals-sources-and-modernization.html

    - **HRSA:** Aims to improve access to healthcare for populations who are uninsured, isolated, or at high risk.
    - **NIH:** Acts as the steward of biomedical and behavioral research across the U.S. and supports public health efforts through the maintenance of health data repositories (e.g., NLM digital sequence information) and public outreach to promote informed health decisions.
    - **SAMHSA:** Leads public health efforts to advance the behavioral health of the nation and improve the lives of individuals living with mental and substance use disorders, as well as their families.
- **The public:** The general population plays a crucial role in public health through participation in preventive measures and other actions, which include vaccination, hygiene, personal health and lifestyle, and disease and symptom reporting. This includes individuals that are beneficiaries of services and their caregivers.
- **Other federal agencies:** Federal agencies external to HHS, such as the Environmental Protection Agency (EPA) and Department of Education (ED), are critical partners and data providers to support public health actions. This includes information provided through public benefit programs, population data, environmental data, and job and economic data.
- **Public Health Service Commissioned Corps:** The cross-agency work and value across federal agencies is exemplified throughout the U.S. Public Health Service Commissioned Corps. Commissioned Corps officers serve 21 federal agencies, demonstrating the important strategic role all agencies have in responsible AI adoption to support public health.
- **STLTs:** STLTs and freely associated state health departments are the backbone of the public health ecosystem and are key partners to HHS in public health work. STLTs are responsible for the health and wellness of their communities and critically manage datasets that are shared with federal health agencies and support prevention and interventions within their communities (e.g., vaccine distribution) as well as issue guidelines to various local stakeholders and organizations.
- **Public education and outreach organizations:** Communication and public education campaigns are a critical component of the public health value chain, including the promotion of immunization campaigns. There are several entities (e.g., the Medicare PACE program) which represent the Surgeon General and the U.S. Public Health Service Commissioned Corps, whose mission is to perform public health education and outreach, such as promoting immunization campaigns. These organizations may receive federal funding tied directly to specific campaigns, funding from private partners, or funding from healthcare organizations in local areas. Providers, payers, and CBOs also play critical roles.
- **Academia and research institutions:** Academic and research organizations, including associated hospitals, labs, and research institutions, are key producers of scientific research. They provide training for the next generation of public health staff and serve as a critical hub for innovation across the field, particularly related to AI use cases.
- **Healthcare systems, providers, and labs:** Healthcare systems are critical for the successful delivery of ongoing and emergency public health programs, and serve as producers of data through health registries, surveillance systems, and research databases, which inform policy.
- **Pharmaceutical, biotechnology, and medical device industry:** Private life sciences organizations support public health efforts through the provision and distribution of drugs, biological products, and medical devices at scale. They also include researchers and subject matter experts involved in medical research and discovery and are major sources of AI innovation.
- **Global partners:** Global partners, including multilateral organizations, bilateral organizations, NGOs, foreign governments, and others collaborate with U.S. public health agencies to address health challenges that transcend borders. Their collective actions help facilitate the sharing of knowledge and data, support the early mitigation of infectious diseases, prevent public health emergencies, support capacity building in healthcare systems, and help ensure equitable access to healthcare services and interventions across regions.
- **Non-profit and CBOs:** National public health collaborative organizations, whose membership typically consists of people and entities dedicated to a particular public health function (e.g., epidemiologists) or purpose (e.g., strengthening public health laboratories) play an important role as partners, conduits, and implementation intermediaries to federal and STLT public health agencies. Additionally, NGOs embedded

in communities support the delivery of health and wellness services to ensure that public health programs reach vulnerable populations effectively.

- **Foundations and private funders:** Foundations may support public health by providing funding for clinical trials and research in areas such as SDOH, or directly delivering public health services. Additionally, other funders may invest in organizations, such as technology companies, in the value chain.
- **Technology companies:** These include companies focused on AI infrastructure (e.g., cloud storage), large, diversified tech companies, vendors of digital solutions, and white hat hackers. These companies provide the infrastructure and services for stakeholders to adopt AI.

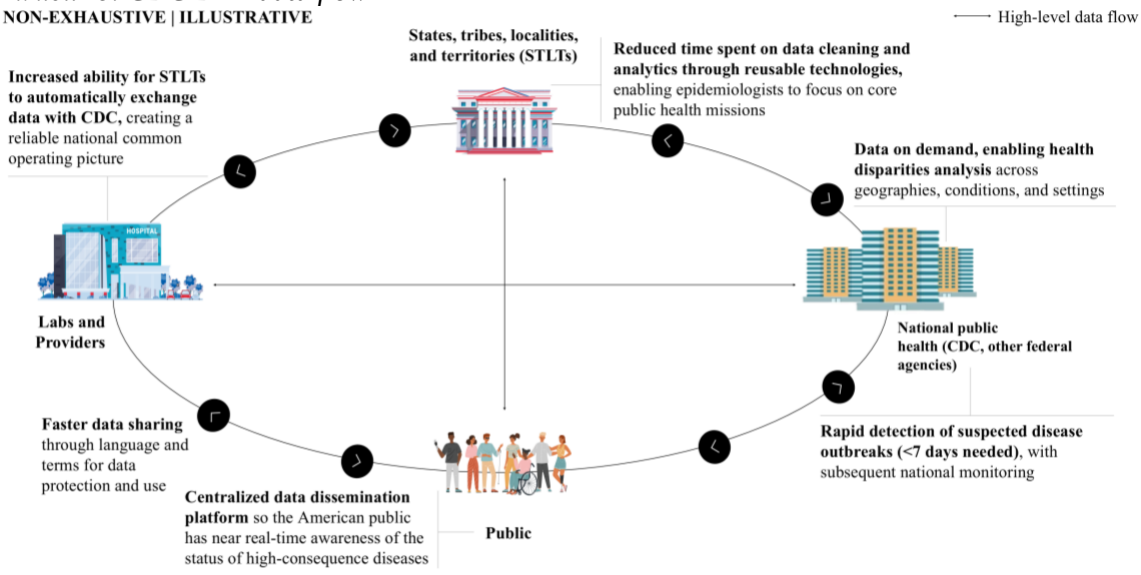
Several HHS divisions (e.g., ASTP, OGA, NIH, and others) advance global health AI efforts through bilateral and multilateral collaboration, conferences, and multi-national organizations, such as the Global Digital Health Partnership, a collaboration between WHO and country governments to support the executive implementation of worldwide digital health services. Additionally subject matter experts from across the Department act as delegates to provide policy input and feedback to multinational organizations such as the Group of Seven (G7), Group of Twenty (G20), and the Organisation for Economic Co-operation and Development (OECD).

The data value chain

The stakeholders above play many pivotal roles in public health, including data collection. As discussed earlier, without high-quality data, proper data collection, and standardization, the ability of AI to drive insights may be limited. The sections below outline the data value chain in public health, existing data improvement efforts from the CDC, and additional actions to strengthen the public health data ecosystem.

As a central player across the public health data flow (Exhibit 13), the CDC has already begun making significant headway through its Data Modernization Initiative (DMI), an effort focused on improving the accessibility, timeliness, and comprehensiveness of data for day-to-day public health responses. The DMI seeks to address public health functions like improved data-sharing speed (e.g., through language and terms for data protection and use), increased ability for STLTs to exchange data with CDC (e.g., through automatic pipelines), and enabling near-real-time public reporting of diseases (e.g., through a centralized data dissemination platform). Many of these investments to create a unified approach to data management at all levels of public health can lay the foundation to support additional AI use cases.

Exhibit 13: CDC DMI data flow[662]



[662] https://www.cdc.gov/ophdst/public-health-data-strategy/public_health_data_strategy-final-p.pdf

Notable recent DMI accomplishments include:

- **Connecting public health and healthcare systems** by aligning current data infrastructure with requirements to exchange information through TEFCA™, supporting adoption of interoperability standards like the USCDI and the USDCI+ initiative, and using intermediaries to reduce point-to-point connections (Exhibit 13). For example, in 2023, CDC helped connect 90% of Epidemiology and Laboratory Capacity recipients to the Association of Public Health Laboratories Informatics Services, ReportStream, or health information exchanges for lab data.[663]
- **Automating and improving data access** by supporting the implementation of automated bidirectional electronic reporting feeds like electronic case reporting (eCR), electronic laboratory reporting, and admission-discharge-transfer feeds to reduce manual reporting. In 2023, the CDC helped 34 jurisdictions implement eCR data to improve case monitoring.[664]
- **Streamlining data collection and processing** to help ensure data is collected once and reused across public health entities, reducing duplication and improving integration. CDC is migrating toward an integrated cloud-computing data platform[665] and using AI use cases for key data systems (e.g., modernizing the National Vital Statistics System to automatically code multiple causes of death).[666]
- **Implementing a core data use agreement (DUA)** to unify and enhance data exchanges nationally across jurisdictions.[667]

Continuing these efforts, with additional integration across public health, healthcare (particularly primary care), and community organizations, could further build resilience and improve healthcare services in both emergency response and everyday contexts (e.g., through automated data exchange across healthcare system EHRs and local demographic data from CBOs). There is a bold opportunity to build on CDC's and others' efforts to further integrate data, which could both be supported by AI and enable AI use to advance public health priorities.

5.3 Opportunities for the Application of AI in Public Health

As AI technologies become more widespread, HHS will work to ensure AI is integrated within public health organizations and missions in an ethical, dependable, and equitable manner by public health partners. There are multiple opportunity areas where AI can support public health priorities and infrastructure:

1. **Improving threat detection, data-driven decision-making, and the effectiveness of interventions:** There is an opportunity to use AI in aggregating and analyzing larger, more complex, or unstructured health datasets, including healthcare delivery data (e.g., claims or EHR data)—in addition to non-health datasets (e.g., migration patterns and climate)—that could support more timely and effective interventions. One example of this is the integration of secondary data into surveillance systems to better predict and respond to emerging public health threats.[668] Another example of this is integrating SDOH and other datasets to better understand underlying risk factors and disease processes for non-communicable diseases, such as diabetes and cardiovascular diseases, to inform effective intervention design. Lastly, a broader application (e.g., data exploration) could be used to accelerate guideline development by supporting the initial synthesis of research across disease areas, provided there is appropriate human oversight and transparency. While these

[663] https://www.cdc.gov/public-health-data-strategy/php/about/phds-progress-in-2023.html

[664] https://www.cdc.gov/public-health-data-strategy/php/about/phds-progress-in-2023.html

[665] https://www.cdc.gov/data-modernization/php/technologies/edav.html

[666] https://www.cdc.gov/surveillance/data-modernization/technologies/ai-ml.html

[667] https://www.cdc.gov/data-interoperability/php/use-agreement/index.html

[668] https://www.ncbi.nlm.nih.gov/books/NBK11770/. As defined by the authors of this book, "Public health surveillance is the ongoing systematic collection, analysis, and interpretation of data, closely integrated with the timely dissemination of these data to those responsible for preventing and controlling disease and injury" and does not constitute any other forms of surveillance.

are just examples, the deployment of AI in public health requires the development and adoption of guidelines to address the associated ethical and safety implications.[669]

2. **Optimizing the allocation of limited resources, especially during public health emergencies:** Resourcing has long been a challenge for the U.S. public health system, where funding may be siloed and sometimes inconsistent.[670] To best prioritize efforts that maximize health outcomes and equity, AI can be used to identify high-risk areas where existing interventions can have the most impact—ensuring the right resources reach the right communities at the right time. During the COVID-19 pandemic, many countries, including the U.S., prioritized vaccine distribution to high-risk populations like frontline workers and the elderly, an effort that was supported, in some cases, by predictive analytics.[671] Emergency response can be further enabled through AI in various ways, such as predictive modeling of supply chains and mapping of vaccine acceptability.[672, 673] Recently, CDC has been using AI to help inform interventions and accelerate responses to outbreaks.[674] AI can support resource optimization, both during public health emergencies and in ongoing programs addressing other areas, such as non-communicable diseases.
3. **Improving efficiency of public health operations and supporting public health workers to better serve their communities:** Responsible, safe, and strategic adoption of AI across the public health ecosystem could greatly reduce the operational burden on a healthcare system and public health authorities that are challenged by burnout and excessive workload (e.g., in 2022, 46% of health workers reported feeling burned out often or very often).[675] For example, AI can be leveraged to automate processes related to grant writing and review to reduce costs or time-consuming activities like data entry and compliance reporting, provided there is sufficient human oversight. The administrative burden in public benefits programs is estimated to range from 15% to 30% of total healthcare spending, half of which includes routine or repetitive tasks that could be automated.[676] Current AI tools are not always fit for purpose for these specific tasks and will require additional development to ensure they balance effectiveness with safety and accuracy. Additionally, in order to use these tools, public health professionals will need upskilling and training opportunities on the safe and effective use of AI.
4. **Enhancing health equity and access to care for underserved populations:** AI applications, GenAI in particular, offer unique potential to transform the way public health decisions and programs are implemented, particularly for traditionally underserved populations, provided potential biases are adequately prevented. AI can advance health equity and improve access to care through the elimination of human bias in decision-making, more targeted outreach (e.g., identification and outreach to high-risk, high-need populations), and evidence-based personalized messaging (e.g., based on language needs, health literacy, and local community context) that can increase public awareness and acceptance of public health guidelines and programs.[677, 678] For example, the CDC launched the Coronavirus Self-Checker Chatbot in 2020 to help individuals decide whether to seek care or manage their symptoms at home.[679] Additionally, through automatic capabilities like translation, transcription, and personalization, AI can rapidly generate content that meets the health literacy, language, and local contexts of diverse populations. AI could also be used to support training and guidelines that support the public health workforce or service recipients.

---

669 http:/dx.doi.org/10.5888/pcd21.240245

670 https://www.milbank.org/quarterly/articles/covid-19-and-underinvestment-in-the-public-health-infrastructure-of-the-united-states/. Maani, et al. "COVID-19 and Underinvestment in the Public Health Infrastructure of the United States," Milbank Quarterly (May 2020)

671 https://pmc.ncbi.nlm.nih.gov/articles/PMC8036633/. Jain et al., "A Rapid Review of COVID-19 Vaccine Prioritization in the US: Alignment between Federal Guidance and State Practice," International Journal of Environmental Research and Public Health," (March 2021)

672 https://www.sciencedirect.com/science/article/abs/pii/S0141813024074518

673 https://www.nature.com/articles/s41598-024-76891-z

674 https://www.cdc.gov/surveillance/data-modernization/technologies/ai-ml.html

675 https://www.cdc.gov/vitalsigns/health-worker-mental-health/index.html

676 https://academic.oup.com/healthaffairsscholar/article/2/2/qxae008/7591560

677 Fisher, S., Rosella, L.C. Priorities for successful use of artificial intelligence by public health organizations: a literature review. *BMC Public Health* **22**, 2146 (2022). https://doi.org/10.1186/s12889-022-14422-z

678 Chen, Y., Clayton, E. W., Novak, L. L., Anders, S., Malin, B. Human-Centered Design to Address Biases in Artificial Intelligence. *J Med Internet Res* **25**, 43251 (2023). https://doi.org/10.2196/43251

679 https://time.com/5807914/cdc-bot-coronavirus/

5.4 Trends in AI in Public Health

Technological and scientific advancements have accelerated public health improvements throughout history, a phenomenon that was brought to the global forefront most recently with the COVID-19 pandemic. The global public health ecosystem delivered a safe and effective COVID-19 vaccine and achieved greater than 70% coverage worldwide by August 2024.[680, 681] Despite this success, the pandemic brought to light the multiple challenges (e.g., outdated systems and limited resources) and opportunities for innovation in the U.S. public health system. AI technologies show promise for mitigating future public health crises and strengthening the public health system. Notable emerging trends include, non-exhaustively:

1. **Enthusiasm and concerns accompany AI adoption in public health:** There is growing excitement about using AI in healthcare and public health, but there are also concerns among the American public and health officials about its potential impacts. Some public health stakeholders are rapidly adopting AI and referencing it frequently, and are excited to continue provided that concerns with respect to equity and ethics are addressed.[682] For instance, as early as February 2022, there were already more than 4,500 scientific papers referenced the use of AI and ML in response to the pandemic, including 239 on surveillance and 219 on forecasting.[683] In contrast, over half of adults in a recent nationwide survey were unsure of the impact of AI on those seeking health information online, and another 23% felt AI was doing more harm than good.[684] Both AI excitement and concerns will need to be addressed in the future.
2. **Predictive analytics are enabling early disease detection and can accelerate public health responses:** The nowcasting and forecasting capabilities of AI are revolutionizing epidemiology by providing real-time surveillance and predictive modeling that inform proactive public health responses.[685] AI extends the library of diverse data types that can be used to predict and track public health threats.[686] Potential examples include image recognition to aid in the early detection of diseases from medical scans, audio analysis for monitoring mental health through voice patterns,[687] and NLP insights from textual health records and social media to track disease spread and public sentiment. There has been strong momentum in the use of AI in these areas.
3. **Implementation of AI in public health is often limited due to resource constraints, infrastructure deficiencies, lack of technological knowledge, and data paucity:** Limited resources in traditionally underserved populations and more fragile environments can contribute to data paucity and difficulty establishing the necessary technical infrastructure that AI relies on. Outdated data infrastructure may also limit the ability to use AI, although initiatives like CDC's DMI and grants to STLTs are helping improve core infrastructure.[688] In addition, public health entities often face challenges attracting, retaining, and training high-quality technical talent; this challenge was exacerbated by the COVID-19 pandemic.[689]
4. **Adoption and use of AI in public health is inconsistent:** Public sector domains, including public health, have unevenly leveraged AI and have varying levels of AI awareness and expertise.[690] Much of this is driven by the availability of high-quality data, differing domain needs, planned and ongoing collaboration efforts, and the availability of modernized data platforms and funding. As discussed above, the diverse potential of AI merits greater investment in widespread implementation.

[680] Vaccine coverage defined as the share of the population that received at least 1 dose of the COVID-19 vaccine.

[681] https://ourworldindata.org/covid-vaccinations. Accessed December 2024.

[682] https://www.healthaffairs.org/doi/10.1377/hlthaff.2024.00050

[683] https://blogs.cdc.gov/genomics/2022/03/01/artificial-intelligence-2/

[684] https://www.kff.org/health-misinformation-and-trust/poll-finding/kff-health-misinformation-tracking-poll-artificial-intelligence-and-health-information/

[685] https://blogs.cdc.gov/genomics/2022/03/01/artificial-intelligence-2/

[686] https://www.cdc.gov/surveillance/data-modernization/technologies/ai-ml.html

[687] https://pmc.ncbi.nlm.nih.gov/articles/PMC11179519/

[688] https://www.cdc.gov/surveillance/surveillance-data-strategies/dmi-investments.html

[689] https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2024.00020

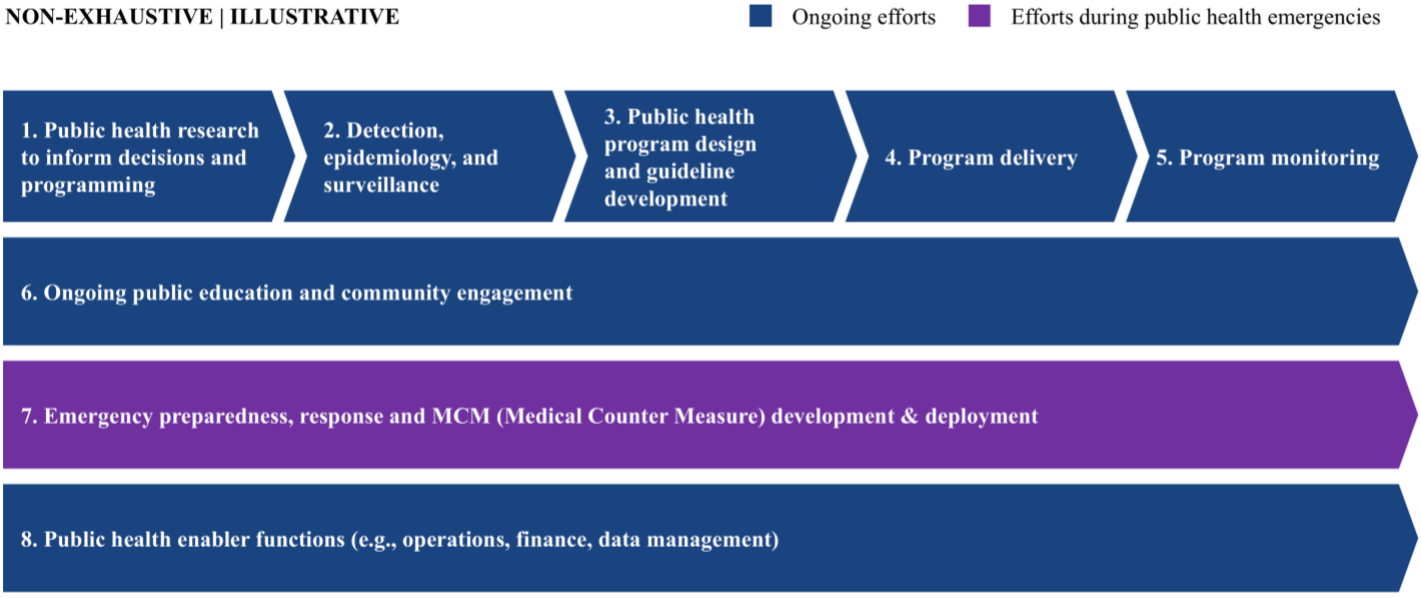
[690] https://www.tandfonline.com/doi/full/10.1080/14719037.2023.2231950

5.5 Potential Use Cases and Risks for AI in Public Health

The below value chain, while non-exhaustive, highlights core public health operations and program areas, with a particular emphasis on preparedness and response during acute public health emergencies. For further details on related topics, in particular refer to the following chapters: Medical Research and Discovery and Medical Product Development, Safety, and Effectiveness for the product development life cycle, which public health informs; Healthcare Delivery for the delivery of healthcare, which is inextricably linked to public health; and Human Services for the delivery of programs that often address SDOH.

This framework is an illustrative representation of the diversity of public health AI applications and should be adapted to the specific contexts that organizations operate in, including areas such as infectious disease control, chronic disease prevention and management, and more.

Exhibit 14: Public Health Value Chain



Every step in the public health value chain represents opportunities for AI to improve the work of public health in the form of increased efficiency, greater analytical power and complexity, improved healthcare and information access, and broader awareness of public health priorities.

At the same time, AI is accompanied by risks, many of which are found across multiple use cases. In public health applications, some common risks include bias (intentional or unintentional discrimination of certain groups due to flaws or underrepresentation in training data),[691] confabulation (fabrication of sources or information),[692] poor interpretability (difficulty explaining AI results due to large datasets and multiple parameters), parasocial relationships (user interpretation of socialization due to "lifelike" interactions with AI models), [693] and unauthorized disclosure of confidential information. Additionally, without comprehensive guardrails, AI may be adopted over traditional techniques for cost savings, even if AI is less effective.

[691] https://www.cdc.gov/pcd/issues/2024/24_0245.htm

[692] https://www.cnn.com/2023/08/29/tech/ai-chatbot-hallucinations/index.html

[693] https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/how-deep-is-ais-love-understanding-relational-ai/77364078496FCE70F71C7A9F293AC322 Gillath, O., Abumusab, S., Ai, T., et al. How deep is AI's love? Understanding relational AI. *Behavioral and Brain Sciences*. 2023

An AI risk that is particularly challenging for public health is misinformation and disinformation; as the COVID-19 pandemic showed, "information that is false, inaccurate, or misleading according to the best available evidence at the time" is now able to spread at never-before-seen speed and scale (e.g., through social media and search engines) and can lead to serious public health consequences like harassment and violence against health workers, insufficient adherence to quarantine guidelines, and the promotion of unproven medical treatments.[694] In addition, given that GenAI is built on neural networks with multitudes of parameters, it is difficult to explain how insights and recommendations are generated—sometimes referred to as the "black box problem." [695] Combined with the potential for false responses and simulated deepfakes (realistic-looking fake images, audio, or video), AI may impact national trust and ultimately reduce the effectiveness of public health programs.

Some of the risks are further considered below; HHS will continue to support mitigation against these risks, in alignment with the action plan discussed later in this document.

In the tables below, HHS highlights a non-exhaustive list of potential benefits and risks of AI across the public health value chain. Please note that the use cases detailed below highlight existing or potential ways that AI can be used by a variety of stakeholders in this domain. For details on how HHS and its divisions are using AI, please reference the HHS AI Use Case Inventory 2024.[696]

Functional component 1: Public health research to inform decisions and programming

Research and analytics efforts to understand and improve the health of populations and inform future programs and decisions

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Deriving novel insights through rapid analysis of large, complex, and often unstructured datasets to inform programming AI can enable greater data capture and analysis of previously unused or underutilized data beyond traditional tabular and numeric formats (e.g., audio files and images) to inform more effective interventions. *E.g., NLP of health records* AI algorithms can extract clinical insights from unstructured EHRs and conduct prediction and classification tasks that would be challenging to do using traditional methods; this can inform public health interventions and population studies.[697]	**Potential to introduce bias and discrimination** *E.g., exclusion of underrepresented groups* AI is often trained on historical data, which often focuses on specific demographic groups more than others, leading to misrepresentative findings that do not apply equally across groups and perpetuation of existing biases.[698, 699]

694 https://www.hhs.gov/surgeongeneral/priorities/health-misinformation/index.html
695 https://doi.org/10.1016/S2589-7500(21)00208-9
696 https://www.healthit.gov/hhs-ai-usecases
697 https://pubmed.ncbi.nlm.nih.gov/36805219/
698 https://postgraduateeducation.hms.harvard.edu/trends-medicine/confronting-mirror-reflecting-our-biases-through-ai-health-care
699 https://pmc.ncbi.nlm.nih.gov/articles/PMC6347576/

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Integrating multiple data types and secondary data in existing public health models to inform research and effective interventions Integrating health datasets (e.g., case data and wastewater data) and non-health data (e.g., migration patterns, travel and sales patterns, and search engine data) can inform forecasting and surveillance of ongoing public health priorities and emerging threats. *E.g., integration of non-health and individual healthcare data with public health data in non-communicable disease and other disease contexts* AI can help generate interpretable insights on how previously unaccounted-for factors (e.g., SDOH, environmental and digital) influence disease risk and analyses that can be further improved by the integration of existing healthcare (e.g., claims data) and public health datasets (e.g., epidemiological data).	**Potential to reduce validity or interpretability** *E.g., unclear conclusions due to multifactorial data* Large datasets that include multiple variables may inaccurately find associations where no true connection exists. **Potential to overuse synthetic data** *E.g., degradation of model integrity and diverse representation as synthetic data is iterated on* Using synthetic data, even with positive intent to increase diversity, can erode model quality as it is analyzed, re-analyzed to produce additional synthetic data, and so on. This could jeopardize the accuracy and validity of results and ultimately not achieve the potential goals of representing diverse populations and/or reducing bias.
Using synthetic data and data linkage techniques to advance research and preserve privacy Synthetic data, which is artificially generated to mimic patient or population data without containing any actual personal information, allows researchers to conduct studies and test algorithms without the risk of exposing personally identifiable information (PII) or PHI. PPRL techniques further enhance this capability as it allows for the matching of records corresponding to the same entity across different databases. *E.g., digital twins and scenario modeling* Virtual replicas of physical systems or processes, like personalized patient models, can be used to simulate different scenarios (e.g., public health scenario modeling and clinical trials) to optimize public health interventions or treatment plans and improve outcomes.[700]	

Functional component 2: Detection, epidemiology, and surveillance

Data and models used to analyze disease trends, identify outbreaks, and study the distribution and determinants of health events in populations

For more information see the Healthcare Delivery and Medical Product Development, Safety, and Effectiveness chapters

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Leveraging AI-powered infectious disease surveillance and prediction AI models can be leveraged to process high volumes of health and secondary non-health datasets to signal potential hotspots or outbreaks as well as monitor ongoing disease spread and changes. E.g., *AI-enabled syndromic surveillance* In parallel with traditional statistical approaches, AI methods can be used to analyze data to detect emerging health threats.[701]	**Potential to reduce interpretability** *E.g., false identification of disease trends* AI models that are overly sensitive to small disturbances may falsely report variations as public health events,

[700] https://www.nature.com/articles/s41746-023-00927-3
[701] https://pmc.ncbi.nlm.nih.gov/articles/PMC7484813/

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Developing intelligent disease diagnostic tools to improve clinical decision-making AI and ML algorithms can be harnessed to improve clinical decision-making and diagnostics from imaging systems to advance detection of non-communicable diseases and ongoing public health priorities (e.g., AI-powered detection of cardiac heart failure). *E.g., AI-powered image processing and diagnostics* AI can be used to accurately analyze images (e.g., mammograms) and diagnose diseases to support clinical decision-making or accelerate public health screening campaign.[702, 703]	inappropriately directing public health efforts. **Potential to disclose confidential information** *E.g., unauthorized disclosure of PHI* Detection algorithms with access to PHI may inadvertently reveal PHI or other identifying information in outputs or be subject to cybersecurity threats.
Advancing precision public health to optimize resources Integrating precision medicine (e.g., genomics and metabolomics) with population-based strategies can help provide "the right intervention to the right population at the right time."[704] *E.g., identification of high-risk geographies or populations* Precision public health can identify vulnerable communities, enabling public health entities to take proactive action.	

[702] https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30160-6/fulltext
[703] https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.122.060137
[704] https://www.ajpmonline.org/article/S0749-3797(15)00522-X/abstract

Functional component 3: Public health program design and guideline development

The creation of strategies and interventions to improve population health outcomes and prevent disease and persistent health issues (e.g., cancer and diabetes). Also includes the development of public health guidelines (e.g., vaccine recommendations and postmarket monitoring of medical products)

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Designing hyper-local public health programming to optimize resources AI can be leveraged to aggregate and analyze local health data to better understand targeted needs (e.g., prevalence of disease by neighborhood), risks (e.g., environmental and socioeconomic factors), and/or infrastructure capacity (e.g., healthcare worker availability, access to PPE, access to testing and access to nutrition) to create targeted programs that optimize resource usage. *E.g., crowd-sourced air quality analytics and advocacy* Using AI to integrate community input and various data sources (e.g., civilian reports and photographs and emission data) enables research that not only advances science but also drives social change (e.g., identification of practical actions with immediate effects based on data on local pollution patterns and their health effects).[705]	**Potential to divulge confidential information** *E.g., unauthorized disclosure of PHI* Algorithms with access to PHI may inadvertently reveal PHI or other identifying information in outputs or be subject to cybersecurity threats. **Potential to design impractical interventions** *E.g., increasingly narrow design not applicable to broad populations* AI models focusing on population subset analysis may narrow program design leading to the creation of highly specific interventions that cannot be scaled across communities, an impractical outcome in public health where issues are widespread.
Automating grant and Request for Proposal (RFP) writing or reviewing processes to improve efficiency Enabled by human oversight and transparency, AI can automate select manual steps in the grant and RFP writing or reviewing processes (e.g., aggregating data, proofreading to ensure accuracy and compliance with submission requirements) enabling substantial efficiencies for government entities and non-profits. *E.g., grant writing assistant apps* Based on user inputs like length of response, conciseness, and context, grant writing assistants can enable supporting initial drafts with appropriate oversight and transparency. *E.g., grant reviewing tools* AI tools can rapidly sort through RFP responses or proposals to synthesize key trends or gaps in the applications, supporting the human-led review process.	**Potential to misunderstand or mischaracterize grant applications** *E.g., federal due process concerns* Mistakes by AI that violate federal regulations governing grant approval or continuation can lead to due process concerns for grant applicants and concerns about the proper allocation of government resources, potentially leading to litigation.

[705] https://airquality.lacity.gov/

Functional component 4: Program delivery

Implementation and administration of public health programs

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Personalizing program delivery to enhance access and equity Public-facing AI tools can be used to efficiently dispense personalized health advice or programming across broad populations and diseases. *E.g., AI chatbots* Chatbot apps and interfaces can conduct conversations and generate a wide range of non-scripted, conversational responses based on user text or voice input (e.g., CDC's COVID-19 chatbot and WHO's S.A.R.A.H GenAI tool delivers tailored messages on well-being topics like nutrition and stress management based on user video or text inputs).[706] *(see Functional component 6: Ongoing public education and community engagement for further details)*	**Potential to confuse users or provide inaccurate recommendations** *E.g., misidentification of AI as human* Users may confuse AI chatbots with human interaction, developing emotional attachments or other parasocial relationships with adverse mental health effects.[707] **Potential to disenfranchise the workforce** *E.g., belief that public health staff are being replaced* Public health experts may perceive the role of AI in program delivery as replacing their roles. **Potential to reduce staff skillset** *E.g., declining skills for community health workers (CHWs) or others* Decreasing interactions between CHWs and the people they serve prevents the close understanding and connection necessary for CHWs to serve as liaisons between health/social services and communities.
Improving program delivery speed and reach AI tools can be used to accelerate program delivery through faster performance of manual tasks and broader reach (e.g., virtually instead of in person). *E.g., food product sampling* While food sampling for safety and quality is traditionally performed manually, AI can be used to automatically conduct sampling with improved accuracy and repeatability. This would enable faster detection of potential outbreaks, reducing the spread of disease.[708] *E.g., supporting public health campaign delivery* Using AI to predict areas or populations in need of additional resourcing given changing factors (e.g., rapid processing of healthcare usage, disease prevalence, and resource availability data) to make dynamic changes to resource prioritization and allocation or campaigns.[709]	

[706] https://www.who.int/campaigns/s-a-r-a-h

[707] https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/how-deep-is-ais-love-understanding-relational-ai/77364078496FCE70F71C7A9F293AC322 Gillath, O., Abumusab, S., Ai, T., et al. How deep is AI's love? Understanding relational AI. *Behavioral and Brain Sciences*. 2023

[708] https://www.fda.gov/food/new-era-smarter-food-safety/new-era-smarter-food-safety-blueprint

[709] https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2780137

Functional component 5: Program monitoring

Tracking and assessing the progress, compliance, and effectiveness of public health programs

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Improving detection of data issues and abnormalities to enhance program effectiveness and efficiency AI-based program monitoring can continuously and proactively identify data anomalies or outliers in data that may indicate potential health issues or errors. *E.g., detection of data abnormalities* AI models can be used to detect unexpected program results, which can facilitate the detection of adverse events (e.g., malfunctioning device) or anomalies (e.g., unusual results, potential fraud) that merit further investigation or adjustment.	**Potential for incomplete or incorrect analysis** *E.g. limited integration of local context or participant feedback* AI-driven program analytics might not appropriately consider qualitative participant feedback or contextual factors that influence program performance (e.g., cultural, economic, or social conditions), leading to less effective decision-making.

Functional component 6: Ongoing public education and community engagement

Two-way, continuous efforts to inform, engage, and collaborate with communities and individuals to improve health education outcomes and program sustainability

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Personalizing public health messaging and education to increase access and improve equity AI can be leveraged to curate messages specifically for different demographics and scale outreach to a broader audience at low cost (e.g., 24/7 availability, automatic content moderation). *E.g., content generation tools* AI tools can assist organizations in tasks like drafting email campaigns, creating appealing webpages, and personalizing content that appeals to specific populations.[710]	**Potential to produce inaccurate messaging** *E.g., delivery of inappropriate messaging to specific populations* AI used for content creation may misinterpret audiences and deliver either inappropriate content (e.g., culturally insensitive recommendations) or inappropriate tone (e.g., medical terminology to the general public). *E.g., spread of misinformation and/or disinformation* Especially in uncertain or evolving situations, misinformation and/or disinformation can be enabled by AI deepfakes and other false content and spread rapidly over the internet, stoking public uncertainty and mistrust of prevailing public health guidelines.
Fostering inclusive communication to improve access and increase equity AI-based translation technologies can help bridge linguistic and / or cultural gaps, enabling organizations to reach a diverse audience at scale. *E.g., real-time translation apps* Advanced translation tools can enable live interactions adapted to specific languages, dialects, or jargon, which can help build trust, advance equity, and increase engagement.[711]	

[710] https://www.cdc.gov/health-communication/media/pdfs/2024/10/AI-for-Good_Listen-Up_S2E5_Transcript.pdf

[711] https://pubmed.ncbi.nlm.nih.gov/37904073/ Bakdash, L., Abid, A., Gourisankar, A., Henry, T. L. Chatting Beyond ChatGPT: Advancing Equity Through AI-Driven Language Interpretation. *J GEN INTERN MED* 39, 492–495 (2024)

Functional component 7: Emergency preparedness, response, and medical countermeasure development and deployment

Design, coordination, and implementation of strategies and interventions to prevent, detect, and respond to acute health threats

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Supporting public health emergency personnel to increase the efficiency and effectiveness of their response During public health emergency and response situations, AI can be used to reduce the immediate burden faced by staff, increase the efficiency of training and onboarding programs (e.g., tailored healthcare worker programs based on local context), and support rapid response (e.g., resource allocation).[712] *E.g., self-learning rescue robots* AI-based robotic systems can offer support in disaster prevention and response (e.g., scouting an unknown situation, identifying hazards, and conducting rescue operations in life-hostile environments).	**Potential to misdirect staff** *E.g., inappropriate directions provided to emergency support staff* AI tools leveraged in emergency situations may misinterpret emergency hazards (e.g., fires and flooding) and recommend actions with the potential to cause harm if inaccurate or misinterpreted.
Disseminating real-time public health guidelines to improve access AI systems can be used to rapidly integrate data sources, generate alerts, and target the distribution of emergency messages. *E.g., weather advisory messages* SAI models can analyze geographic, weather, and user data to generate relevant and informative alerts (e.g., different winter weather advisories depending on location and whether the user is driving).	
Developing medical countermeasures AI has the potential to empower more informed decisions on where investments should be directed (e.g., molecules or vaccine R&D pipeline), along with better monitoring of medical countermeasure effectiveness in real-time. *E.g., identification of potential drug targets* These target discovery methods can help uncover novel targets and pathways underlying diseases, enabling faster development of interventions.[713] *Note: Medical countermeasure development is particularly cross-cutting with life sciences and is primarily discussed within the Medical Research and Discovery and Medical Product Development, Safety, and Effectiveness chapters of this Strategic Plan.*	**Potential to misdirect resources** *E.g., inappropriate identification of drug targets* Inaccurate conclusions drawn by AI models (e.g., ineffective drug targets) may falsely build confidence in interventions.

[712] https://www.noaa.gov/news-release/biden-harris-administration-invests-250k-to-develop-powerful-artificial-intelligence-tool
[713] https://www.fda.gov/media/167973/download

Potential use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Evaluating and learning from past emergency response efforts to improve the effectiveness of interventions AI tools can be harnessed to analyze data from past public health responses to identify trends, successes, and opportunities for improvement (e.g., the impact of various state-specific COVID-19 policies to inform future public health decisions). *E.g., AI-enabled after-action reviews* An AI-enabled review process can analyze an organization's response to an emergency or disaster, compare it to a vast library of previous records to identify areas for improvement and develop targeted recommendations and programs (e.g., simulations).	

Functional component 8: Public health enabler functions (e.g., operations, finance, IT, and data)

Infrastructure and administrative processes necessary to support public health service delivery and management

Potential benefits and example use cases (non-exhaustive)	**Potential risks** (non-exhaustive)
Automating administrative and operational tasks to improve efficiency Like many other organizations, AI offers significant opportunities for public health agencies to achieve operational efficiencies and reduce human errors, particularly in the realm of labor-intensive administrative tasks, when combined with human oversight.[714] *E.g., data entry* Smarter data entry facilitated by AI can help transcribe information from various formats into centralized databases and enhance the overall quality of data with predictive text fields and real-time error-checking algorithms.[715]	**Potential to disenfranchise workforce** *E.g., belief that administrative and operational staff are being replaced* Individuals whose roles involve tasks that can be automated may perceive the role of AI as replacing their positions and responsibilities.

5.6 Action Plan

In light of the evolving AI landscape in public health, HHS has taken multiple steps to launch ecosystem-wide infrastructure updates and create guidelines that promote responsible AI. The Action Plan below follows the four goals that support HHS's AI strategy: 1. catalyzing health AI innovation and adoption; 2. promoting trustworthy AI development and ethical and responsible use; 3. democratizing AI technologies and resources; and 4. cultivating AI-empowered workforces and organization cultures. For each goal, the Action Plan provides context, an overview of HHS and relevant other federal actions to date, and specific near- and long-term priorities HHS will take. HHS recognizes that this Action Plan will require revisions over time as technologies evolve and is committed to providing structure and flexibility to ensure longstanding impact

5.6.1 Catalyze Health AI Innovation and Adoption

The adoption and implementation of AI have the potential to revolutionize public health and protect the public against emerging and ongoing threats, for example, through enhanced disease forecasting. Unlike healthcare

[714] https://www.science.org/doi/10.1126/science.adh2586
[715] https://www.healthit.gov/hhs-ai-usecases/ai-assisted-data-entry

delivery or R&D, where the private sector is heavily involved in AI innovation and investment, private sector engagement in public health is more limited. Therefore, HHS will play an even more crucial role in allocating resources, aligning incentives, and guiding AI implementation and adoption across the public health ecosystem.

As such, HHS can address current challenges and barriers to innovation through:

1. Encouraging research, development of guidelines, and identification of resources to support evidence generation and scale of AI in public health
2. Modernizing infrastructure necessary to implement AI and support adoption

Below, we discuss context, HHS actions to date, and plans to catalyze health AI innovation and adoption.

1. Encouraging research, development of guidelines, and identification of resources to support evidence generation and scale of AI in public health

Context:

As AI advances, its full impact remains uncertain, highlighting the need for cross-disciplinary research to encourage widespread innovation and adoption with responsible use. As such, targeted research on the potential of AI for impact across core public health objectives (e.g., health equity, patient privacy) and diverse public health domains (e.g., immunization outreach, emerging disease research) can provide evidence of the effectiveness and cross-domain applicability of AI. HHS and its divisions can lead by example by identifying and prioritizing scalable high-impact AI use cases that address the most pressing public health challenges, from improving disease surveillance and emergency response to addressing limited resourcing and workforce shortages, to advancing health equity and access to care. HHS will continue to create programs, guidelines, and resources to support AI innovation, and share its findings with the broader public health ecosystem to encourage further innovation.

HHS actions to date (non-exhaustive):

- **CDC Data Modernization Initiative (DMI)** is investing in tools and technologies (e.g., advanced disease surveillance systems, real-time data analytics platforms) to get better, faster, actionable insights for decision-making at all levels of public health (see above for additional details).[716]
- **CDC AI Accelerator Initiative (AIX)** focused on operationalizing and scaling four high-impact public health use cases.
- **CDC staff chatbot** is an internal AI chatbot to provide guidelines on interacting with GenAI, enabling staff to innovate safely and responsibly.
- **Public Health Data Strategy AI plan milestone 2.05** outlined a plan for how the agency will leverage AI and launch pilots. CDC hopes to encourage safe and responsible AI use and improve public health efficiency, response readiness, and outcomes through the completion of this milestone.[717]
- **NIH grants and other resourcing programs** like NSF 23-610: National AI Research Institutes or NIH's Bridge2AI provided resources to advance AI use in biomedical and scientific applications.[718, 719]
- **FDA explored the use of AI internally,** including but not limited to: deduplicating non-public adverse event data in the **FAERS**; identifying novel terms for opioid-related drugs using the **Term Identification and Novel Synthetic Opioid Detection and Evaluation Analytics tool**, which uses publicly available social media and forensic chemistry data to identify novel referents to drug products

[716] https://www.cdc.gov/surveillance/data-modernization/index.html
[717] https://www.cdc.gov/public-health-data-strategy/php/about/milestones-for-2024-and-2025.html
[718] https://new.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes/nsf23-610/solicitation
[719] https://commonfund.nih.gov/bridge2ai

in social media text; and searching and indexing tobacco authorization applications using **ASSIST4Tobacco**, an AI-based NLP tool. [720, 721]

HHS near-term priorities:

- Update the **Public Health Data Strategy** to explicitly support AI development and life cycle management.[722]
- Ensure grants related to research through NIH continue to allow for the use of AI and the study of its impacts on public health domains.
- Continue piloting the use of AI to enhance the forecasting of contagious outbreaks, chronic conditions, and addictive substances.
- Continue piloting the use of AI for evidence-based public health messaging to providers and patients tailored to language, literacy, and local context.
- Develop implementation guidelines and playbooks for public health partners on the use of AI models and AI systems used by public health officials to support existing operations using tools commonly available within their systems.
- Partner with nonprofit organizations and others to use AI in health outreach campaigns.

HHS long-term priorities:

- Share findings and impacts of AI on public health, including operational impacts, internal risks, benefits, and other findings to inform future actions and support the broader community.
- Support funding and grants for AI use in public health through existing mechanisms and new opportunities where applicable.
- Consider conducting a strategic review and supporting the scaling up of high-impact investments aligned with division goals. Also, support the alignment of public health partners in these areas.
- Consider supporting guidelines to other stakeholders on how and where to scale and where there may be an investment case.

2. Modernizing infrastructure necessary to implement AI and support adoption

Context:

Many public health entities lack the modern technology infrastructure needed to support AI implementation. As discussed previously, effective public health action relies on integrating diverse data sources (e.g., through public-private data sharing and linkage of existing individual health data with public health data) to enable more holistic patient care. However, current public health data systems are siloed, vary in modernization, and often run on outdated technology, leading to different levels of AI readiness.[723] Additionally, the diversity of data formats and the multitude of data standards limits interoperability and seamless data sharing—for example, CDC currently maintains over 100 separate disease surveillance systems that are not fully integrated.[724] Public health officials may be hesitant to adopt AI solutions due to these technological and resource challenges, which affect the entire public ecosystem's ability to function and communicate effectively. HHS, including CDC and others, has started to lay the foundation for modernizing data systems (e.g., DMI) and is investing significant resources today. However, there are additional actions HHS can and will continue to take.

[720] https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-inventory.pdf
[721] https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-2023-public-inventory.csv
[722] https://www.cdc.gov/public-health-data-strategy/php/index.html
[723] https://jamanetwork.com/journals/jama/fullarticle/2782635
[724] https://pmc.ncbi.nlm.nih.gov/articles/PMC10126962/

HHS actions to date (non-exhaustive):

- **CDC DMI** provided direct funding and technical assistance to STLTs to support eCR (automated data feed) implementation, modernize data infrastructure, and connect public health data systems, among other things
- **Public health infrastructure grants,** as of September 2024, had allocated or distributed $611M in funding to support public health data modernization.[725] This is part of the $4.2B public health infrastructure grant awarded to health departments around the country to support their most pressing needs, from workforce development to laboratory information systems.
- **CDC and ASTP federal interoperability initiative** established TEFCA™, adopted FHIR-based standards for implementing API in certain certified health IT applications and USCDI+ data elements.

HHS near-term priorities:

- Advance HHS Data Strategy to enable cross-agency data sharing to support AI development for public health.
- Pilot use of AI to assist integration and mapping of heterogeneous structured and unstructured public health data streams and public-health-relevant data (e.g., environmental, social media, retail, and over-the-counter medication sales).
- Pilot aggregation of multijurisdictional data for AI development, validation, and risk monitoring.
- Create a strategy for developing AI to support the integration of public health functions into EHR systems.
- Convene regular forums for public health partners to collaborate on data modernization efforts.

HHS long-term priorities:

- Provide additional opportunities based upon available funding and support for grants for data modernization and AI-readiness initiatives.
- Continue implementing data standards across the core public health data systems (e.g., expand the use of USCDI+ data elements and standardize definitions of common data metrics/variables such as population) to improve the quality and completeness of data and maximize AI accuracy and effectiveness.
- Continue current efforts to simplify the technology landscape and help public health entities better integrate and process data (e.g., by implementing key integrated enterprise-wide data platforms of CDC and helping jurisdictions migrate and onboard cloud-based solutions).
- Continue working toward interoperability standards so that AI data systems "speak the same language," including standardized implementation of TEFCA™.
- Continue funding for internal operational capabilities and data modernization for existing core data systems such as Vital Records to increase the processing speed and insights such as identifying trends in opioid-related deaths, drug overdoses, and other pathways.
- Consider additional ways to integrate public health and healthcare data systems or provide opportunities with "sandboxes" for piloting.
- Strategically explore additional ways that AI can both improve the current modernization efforts and where the current modernization efforts could be used as a platform to encourage AI tool use by others.

5.6.2 Promote Trustworthy AI Development and Ethical and Responsible Use

Context: AI has transformative potential to change the way the public, patients, and providers interact with the healthcare system. This includes increasing the tailoring of health information by language, geography, and

[725] https://www.cdc.gov/infrastructure-phig/php/data-research/profiles/index.html

background through the use of AI chatbots, AI-enabled translation tools, and other services.[726] However, for these technologies to be powerful, stakeholders will need to be strongly convinced of their power, trust in the way their data is managed, and be educated on best practices for use. HHS will continue to aim to support innovation in AI use while ensuring safety and privacy.

There are several areas where HHS can have an outsize impact to enable responsible AI use, including:

1. Establishing guardrails to help ensure data quality and accuracy
2. Standardizing data security policies across the public health ecosystem
3. Advancing AI tools and techniques that consider and assess health equity from end to end

Below, we discuss context, HHS actions to date, and plans to promote trustworthy AI development and ethical and responsible use.

1. Establishing guardrails to help ensure data quality and accuracy

Context:

AI models can face issues with data quality and precision, particularly in public health, where inaccuracies can endanger individuals and communities. Both models developed by public health entities and those without public health expertise (e.g., some technology companies) must be trained on appropriate data and parameters to be accurate, reliable, and able to resist misuse to be widely trusted. All models can benefit from robust safeguards to ensure quality and control shared information, such as filters removing explicit content and verifying health data, and continuous monitoring to detect anomalies in real-time. Organizations can also inadvertently or maliciously create biased models or use incorrect data to spread misinformation and mistrust, negative outcomes which are difficult to identify and mitigate. Outside the U.S., there has been recent guidance from public health institutions including the World Health Organization's (WHO) report, and the European Union's AI Act to address some of these challenges.[727, 728] HHS is actively exploring this area and will continue to develop mechanisms, build consensus, and support partnerships to establish and monitor AI standards in collaboration with other authorities.

[726] https://pmc.ncbi.nlm.nih.gov/articles/PMC10637620/

[727] https://www.who.int/publications/i/item/9789240029200

[728] https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

HHS actions to date (non-exhaustive):

- **CDC AI Use Guidelines** laid out principles and practices for responsible use, development, and procurement of GenAI use in early 2024, including for public health contexts.
- **HHS Plan for the Responsible Use of AI in Public Benefits** outlined responsible use of AI in automated and algorithmic systems by STLTs in the administration of public benefits such as health screenings.[729]
- **NIH Office of Extramural Research published NOT-OD-23-149, "The Use of Generative Artificial Intelligence Technologies is Prohibited for the NIH Peer Review Process" in June 2023,**[730] which prohibited NIH scientific peer reviewers from using NLP, LLMs, or other GenAI technologies to analyze or formulate peer review critiques for grant applications and R&D contract proposals.
- **CDC Morbidity and Mortality Weekly Report Instructions for Authors (MMWR)** published in June 2023 provided guidelines on AI use in research reporting, including in healthcare and public health contexts.[731]

HHS near-term priorities:

- Promote transparency on the use of data and AI for public health to combat public mistrust in key areas, including the use of data for disease detection and surveillance and the spread of medical misinformation and disinformation.
- Promote innovation sharing and dissemination of best practices through publications on AI-system information, model cards, training information, and open-source system publications.
- Support the safe and responsible use of GenAI with plain language public health outreach and communication efforts such as CDC's **Clean Slate project**, which can highlight the risks of improper usage and outline best practices.
- Develop standards and guidelines on transparency for scientific research and public health communication on the role AI systems will play in its adoption.
- Develop standards and guidelines to ensure public health providers comply with existing federal civil rights laws when using AI.

HHS long-term priorities:

- Design a mechanism to partner with AI-system designers to ensure pre-training of AI models is not based on medical misinformation or disinformation that could threaten public health. This includes ensuring AI-system outputs include the appropriate context and information to share with any medical information. This could be accomplished through partnership with AI-training organizations in the private sector to support broader adoption.
- Continue to partner with organizations to identify and mitigate misinformation in public health; support collaborative partnerships where appropriate.
- Consider ways to implement continuous monitoring and evaluation of AI applications to detect and address potential issues (e.g., models created by malicious actors), partnering with other organizations where appropriate.
- Update and monitor existing public health data and AI governance structures and guidelines applicable across the public health data ecosystem based upon new capabilities, federal AI policy, and STLT AI policy.

[729] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf
[730] https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-149.html
[731] https://www.cdc.gov/mmwr/author_guide.html

2. Standardizing data security policies across the public health ecosystem

Context:

Many existing data policies and guidelines established at the federal and STLT levels were not originally developed with AI technologies in mind. As AI models are often trained or weighted using PII or PHI data, the AI use without sufficient data protection and security policies can pose significant risks to patient privacy and safety. The lack of standardized policies across the ecosystem can also lead to inconsistencies in how sensitive data is managed and protected from entity to entity, increasing the potential for data breaches and potentially leading to reputational loss and legal or financial consequences. Additional guidelines could build on existing HHS work, such as the HHS Cybersecurity Program or the HIPAA Security Rule and be tailored for use in public health.[732, 733]

HHS actions to date (non-exhaustive):

- **HHS common DUA structure policy** supported securely and ethically sharing data from HHS to federal agencies or external organizations.[734]
- **HHS Healthcare and Public Health (HPH) Cybersecurity Goals** included best practices for healthcare organizations and healthcare delivery organizations.

HHS near-term priorities:

- Create ethical guidelines for AI use in the public health ecosystem to help safeguard individual rights and safety.
- Promote guidelines on secure open-source software and data security practices in AI systems within the public health ecosystem.

HHS long-term priorities:

- Continue existing efforts to modernize data infrastructure, including the standardization of core public health data sources and increased privacy protection of individual data through security measures (e.g., implementation of PPRL and PII reduction technologies to prevent the sharing of sensitive information).

3. Advancing AI tools and techniques that consider and assess health equity from end to end

Context:

AI has the potential to advance health equity by improving healthcare provision, mitigating bias in human decisions, and identifying changeable root "drivers" (e.g., neighborhood conditions) that influence health outcomes rather than relying only on demographic data like race and gender.[735] However, it is crucial to continually investigate and address ways in which AI may inadvertently introduce or amplify health disparities (e.g., biases in data can lead to skewed algorithms that disproportionately affect certain populations).[736] Particularly in underserved communities, where prior incidents or improper data usage may have already eroded trust, there may be skepticism regarding the development and use of AI.[737] HHS will continue to strive to promote the use of AI in a manner that advances health equity.

---

732 https://www.hhs.gov/about/agencies/asa/ocio/cybersecurity/information-security-privacy-program/index.html

733 https://www.hhs.gov/hipaa/for-professionals/security/index.html

734 https://www.hhs.gov/web/governance/digital-strategy/it-policy-archive/hhs-policy-common-data-use-agreement-structure-repository.html

735 https:/www.cdc.gov/health-equity/core/index.html.

736 https://www.science.org/doi/10.1126/science.aax2342. Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453

737 https://www.healthaffairs.org/doi/10.1377/hlthaff.2021.01466

HHS actions to date (non-exhaustive):

- **CDC-Georgia Tech Research Institute (GTRI) partnership** convenes CDC's Office of Science and experts from the GTRI to develop guidelines and training resources for public health researchers to navigate health equity challenges related to AI use.[738]
- **NIH's AIM-AHEAD Program** sought to develop a diverse workforce of researchers proficient in AI and address unmet needs in underrepresented communities.[739]
- *For more information, see **CDC AI Use Guidelines** above.*

HHS near-term priorities:

- Gather resources and conduct an educational public event to share mitigation actions against potential harm associated with synthetic AI-generated content intended to defraud at-risk populations of resources.
- Develop model card and system card standards for public health partners and external partners to use for documenting AI systems, including key fields such as intended use, known limitations, potential model biases, and others based upon industry best practices.[740, 741]

HHS long-term priorities:

- Develop guidelines and best practices in conjunction with partners in the different domains of public health to protect health, save lives, and mitigate harms caused by AI specific to each domain.
- In coordination with the appropriate entities, develop and implement education campaigns and outreach efforts to educate at-risk populations on the potential harms of deepfakes and AI-associated misinformation campaigns to public health (e.g., breaching health data through impersonation of providers, disseminating false images and video that appear to be from a trustworthy public health entity).
- Conduct public education and community engagement on AI, which includes actively involving families, communities, and other stakeholders in the development and implementation of public health events. This includes providing resources contingent on the level of need within communities and fostering a two-way relationship that builds trust, shares power and collaborates to support all parties involved.

5.6.3 Democratize AI Technologies and Resources:

Context: AI represents an outsized opportunity for underserved populations and under-resourced healthcare systems and agencies, as it can help improve cost structures, address resource and staffing gaps, and improve overall resource allocation and use. More so than in fields like human services, global data sharing is essential for public health. Disease knows no borders; only with transparent communication and collaboration can outbreaks and pathogens be rapidly identified and contained. Equitable access to AI can yield substantial benefits and a high return on investment, amplifying its impact across multiple domains. HHS can address current challenges through:

1. Creating an environment that enables data sharing across the public health ecosystem
2. Supporting AI adoption, development, and collaboration, especially for STLTs and community organizations who may have limited resources
3. Developing user-friendly, customizable, and open-source AI tools to broaden access and accommodate a diversity of users

Below, we discuss context, HHS actions to date, and plans to democratize AI technologies and resources.

[738] https://www.cdc.gov/surveillance/data-modernization/snapshot/2022-snapshot/stories/ai-impact-health-equity.html
[739] https://datascience.nih.gov/artificial-intelligence/aim-ahead
[740] https://arxiv.org/abs/1810.03993
[741] https://www.xd.gov/blog/creating-a-client-side-model-card-generator/

1. **Creating an environment that enables data sharing across the public health ecosystem**

Context:

As of 2023, CDC maintains a highly complex data infrastructure with over 1,000 data systems, increasing the challenges related to the modernization of capabilities, implementing AI infrastructure, and ensuring minimum data entry. This figure does not include local systems owned and operated by STLT agencies, which are critical to conducting on-the-ground public health activities and conducting outbreak response. State and local public health officials collect and analyze data, make recommendations to local and state leaders based on these data, and aggregate this data to aid in federal decision-making. Effective data-sharing agreements can enable swift, accurate, bidirectional data sharing across the ecosystem, from STLTs and community organizations to federal agencies, enabling all parties to have a reliable understanding of the current state of health across various parts of the nation. HHS will continue to support effective data sharing that can also support AI use.

HHS actions to date (non-exhaustive):

- **In 2023, ASTP published TEFCA™**, a nationwide framework for health data exchange managed by ONC, to help create a reliable national common operating framework. Over 50 public health jurisdictions across the country use TEFCA™ exchange to support eCR.
- **USCDI+ for Public Health** is a collaboration between CDC and ASTP to develop standardized public health data elements building on USCDI.
- **ASTP's HTI-2 Proposed Rule** included the adoption of a Public Health API and other public health-focused capabilities as certification criteria to which EHR could be certified. Additionally, HTI-2 Proposed Rule proposes to expand ONC Health IT Certification Program certification criteria to include criteria applicable to public health IT systems.

HHS near-term priorities:

- Continue federal support of TEFCA™ framework for health data exchange to streamline public health information sharing between healthcare delivery and public health agencies and between public health agencies.
- Continue USCDI+ for Public Health initiatives to enhance nationwide public health data standards.

HHS long-term priorities:

- Coordinate with standards development organizations on standards for AI technologies in public health.

2. **Supporting AI adoption, development, and collaboration, particularly among STLTs and community organizations who may have limited resources**

Context:

Currently, the creation of AI tools can require significant capital, data, and technical expertise, all of which can present barriers to entry that limit AI providers primarily to the private sector or academia.[742] Federal funding for data modernization and supporting systems, prompted in response to the COVID-19 pandemic, has enabled organizations to begin updating data systems, enhance efficiencies in existing systems, and streamline operations.[743] However, these tasks take time and significant investment, resources which are more readily available in the private sector.[744, 745] In contrast, public health stakeholders, especially STLTs and non-profit organizations, dedicate most of their resources to maintaining essential operations and activities. They

[742] https://www.omfif.org/2024/07/how-the-global-south-may-pay-the-cost-of-ai-development/

[743] https://www.cdc.gov/budget/fact-sheets/covid-19/index.html

[744] https://ourworldindata.org/grapher/private-investment-in-artificial-intelligence

[745] https://hbr.org/2022/12/what-companies-need-to-know-before-investing-in-ai

may have limited ability to invest in long-term needs like AI integration and may have a low-risk appetite due to potential negative impacts. HHS has the scale and ability to support AI adoption in smaller jurisdictions and organizations through resourcing, the creation of shared centralized systems and standards, and strategic advice on how to encourage innovation and AI use.

HHS actions to date (non-exhaustive):

- **HHS Plan for Responsible Use of AI in Public Benefits** outlined additional areas of support for STLTs pertaining to promoting AI use in public benefits, including providing information on funding available to STLTs.[746]
- *For more information, see **CDC's DMI above.***

HHS near-term priorities:

- Develop enterprise communication systems with AI-augmented capabilities for local organizations to use to support public health outreach campaigns.
- Develop a plan for providing tools, appropriately controlled data, sandboxes, and infrastructure to STLTs for AI development and experimentation leveraging the **CDC One Common Data Platform.**
- Convene public health communities of practice with STLTs to identify opportunities, surface enablers, and barriers, identify opportunities for knowledge and resource sharing, and share best practices and lessons learned (e.g., through a professional association).
- Share tactical guidelines on how STLTs and community organizations can engage in low-cost, low-risk "safe innovation" (e.g., suggestions on how to set up an AI working group of existing staff and test simple AI use cases that leverage existing or easy-to-access technology and data).
- Encourage and provide guidelines for STLTs to use existing data platforms and available AI systems and tools whenever possible.

HHS long-term priorities:

- Continue initiatives to develop internal operational capabilities and modernize existing core data systems such as **Vital Records**, including developing and maturing associated AI infrastructure capabilities. This investment can improve processing speeds and provide insights, such as identifying trends and disease pathways in opioid-related deaths and drug overdoses.
- Support enhanced system capabilities across the vital statistics operation chain to enhance insights and NLP capabilities with open-text fields and International Classification of Diseases, 11th Revision (ICD-11) collaboration.
- Provide additional opportunities, based on available funding and support, for grants for data modernization and AI-readiness initiatives.
- Continue the implementation of data standards across the core public health data systems, especially in STLTs and community organizations (e.g., expand the use of **USCDI+** and standardize definitions of common data metrics/variables, such as population).
- Continue working toward ecosystem wide interoperability standards so that data systems "speak the same language," including the standardized implementation of AI.
- Implement a series of high-value, scalable AI projects aimed at improving specific domains of public health. These projects aim to provide immediate, real impact previously unattainable without AI technologies, augmenting efforts to solve existing and emerging public health problems (e.g., using AI to identify cooling towers from satellite images can help better direct response efforts during Legionnaires' disease outbreaks).[747]

[746] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

[747] https://www.thelancet.com/journals/landig/article/PIIS2589-7500(24)00094-3/fulltext

3. **Developing user-friendly, customizable, and open-source AI tools to broaden access and accommodate a diversity of users**

Context:

The use of AI in diverse public health settings, especially under-resourced settings, requires the customizability of AI models and increased access to technology like high-speed internet and intuitively designed AI tools. Increasing the availability of low-code or no-code AI platforms, available to the public at low cost, could enable health entities like STLTs and community organizations to develop sophisticated models that meet their communities' unique needs. Recent resources for AI in public health include ASTP's LEAP in Health IT, which provides funding to address emerging challenges that inhibit the development, use and/or advancement of well-designed, interoperable health IT.[748] Going forward, HHS can consider advancing these and other efforts to support the development of open-source AI tools, particularly where they could be most impactful and where there could be shared platforms.

HHS actions to date (non-exhaustive):

- **ASTP's LEAP in Health IT** provides funding for health IT innovations that further the development, use, and/or advancement of well-designed, interoperable health IT.
- **CDC's AI Acceleration Initiative (AIX)** is developing high-impact public health AI pilots focused on tools that are both broadly reusable and address common public health challenges.

HHS near-term priorities:

- Encourage STLTs to utilize existing data platforms and open-source AI systems available through local government programs. By leveraging state data platforms for AI access, STLTs can reduce maintenance costs and enhance AI capabilities across their partners, who may have varying levels of expertise.

HHS long-term priorities:

- Develop shared analytic zones and tools to promote high-value AI use cases across federal and STLT public health partners; identify common public health challenges and data platforms where this approach could have the greatest impact.
- Implement scalable GenAI-powered chatbots and make them easily available and modifiable to STLTs and other public health partners.
- Develop and acquire open-source AI-powered, along with accompanying training materials, to augment existing public health operations and workforce capabilities.

5.6.4 Cultivate AI-Empowered Workforces and Organization Cultures

Context:

Public health departments, though critical for community health awareness, prevention, and interventions, often struggle with resource limitations. AI could reduce the burden on the public health workforce provided integration is mindful of community needs. To integrate AI in public health operations and foster a learning and innovative environment while addressing community needs, HHS can support the development of use cases, training programs, and pipelines, both formal and informal, that equip public health workers with the skills needed to effectively use AI tools.

HHS can address current challenges by:

[748] https://www.healthit.gov/topic/onc-funding-opportunities/leading-edge-acceleration-projects-leap-health-information

1. Augmenting and supporting the public health workforce to address burnout and attrition while improving efficiency and productivity
2. Promoting AI education and community-based AI approaches tailored to each community's unique needs

Below, we discuss context, HHS actions to date, and plans for AI-empowered workforces and organization cultures.

1. Augmenting and supporting the public health workforce to address burnout and attrition while improving efficiency and productivity

Context:

As previously discussed, while the public health workforce faced challenges prior to COVID-19, the COVID-19 pandemic exacerbated workforce issues, accelerating burnout and attrition.[749] One of the greatest concerns about AI adoption is its potential to replace or reduce existing jobs and workers; however, public health currently faces a severe workforce shortage.[750] Nearly half of all state and local public health professionals left their positions between 2017 and 2021, an attrition rate that, if it continues, could leave the public unprepared for future outbreaks and health threats.[751] Although AI cannot replace the cross-jurisdictional and cross-functional collaboration central to public health knowledge sharing and disease response, there is enormous potential to use it to improve efficiency and support the understaffed public health workforce. For example, AI can automate time-consuming or repetitive tasks, allowing workers to focus on more strategic or person-centered work. At the same time, a "human in the loop" approach can ensure oversight and intervention should errors occur.[752, 753] HHS can continue to identify and develop AI use cases that will streamline processes and boost the productivity of existing public health workers. This not only helps alleviate burnout but also encourages further understanding and adoption of AI across the public health ecosystem.

HHS actions –to date (non-exhaustive):

- **CMS AI Playbook** included educational materials that define AI use cases and trends within healthcare delivery, along with applications CMS is currently and is considering using within its own operations and their potential impact on patient care (e.g., wearables, digital twins and customer support).[754]

HHS near-term priorities:

- Create GenAI tools with image/audio editing functions to augment staff capabilities for education and outreach efforts.

HHS long-term priorities:

- Identify opportunities where AI can improve efficiency by automating routine and repetitive tasks like reporting and data entry.
- Invest in training and change-management initiatives to improve AI adoption, highlighting the impact AI can have on improving workforce efficiency and health outcomes, especially with respect to automating routine and time-consuming tasks.
- Consider reviewing holistically the potential impact of AI on the workforce and ways operations may shift within public health (e.g., impact on staff's sense of connection and purpose).

[749] https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2024.00020
[750] https://www.bbc.com/worklife/article/20230418-ai-anxiety-artificial-intelligence-replace-jobs
[751] https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2022.01251
[752] https://cloud.google.com/discover/human-in-the-loop#benefits-of-human-in-the-loop-hitl
[753] https://doi.org/10.1007/s10462-022-10246-w
[754] https://ai.cms.gov/assets/CMS_AI_Playbook.pdf

2. **Promoting AI education and community-based AI approaches tailored to each community's unique needs**

Context:

Community-based and human-centered approaches are widely used in public health, where community members are engaged from research question selection to program delivery and invited to use their lived experiences to identify and implement appropriate interventions.[755] These programs are better able to address the underlying risk factors that cause health issues, empower community members and increase program engagement, and can often reduce the cost of care through multifactorial approaches that address non-medical challenges like food insecurity and lack of transportation.[756] Alongside system upgrades and funding programs, an AI-empowered workforce that understands how and when to use AI (and when not to) and how to engage the community will be needed to ensure AI is used responsibly and effectively.

HHS actions to date (non-exhaustive):

- *See information on **NIH's AIM-AHEAD Program** above*

HHS near-term priorities:

- Define HHS's strategic priorities for promoting awareness and building trust in public health AI.
- Coordinate with academia and schools of public health to ensure students gain skills in implementing responsible and ethical AI efforts through their coursework, degree programs, and other education opportunities.
- Partner with public health collaboratives and professional organizations to integrate core AI skills into communications, competencies, and certifications.
- Develop AI programs and tools that use a community needs approach to incorporate community voices throughout the public health program design and implementation process.

HHS long-term priorities:

- Expand existing education pathways to include opportunities for STLT and federal staff to upskill in operational AI and advanced data science capabilities.

5.7 Conclusion

AI technologies offer a unique opportunity to accelerate the operational efficiencies of public health agencies, advance data gathering, forecasting, and analytics, and improve outreach and communication efforts in a manner that advances equity and improves health outcomes. However, with the many benefits of AI adoption come risks like the potential for AI-enabled misinformation campaigns through deepfakes sharing harmful health advice. Over the coming years, HHS can build upon the foundation of data modernization and innovation laid through the COVID-19 pandemic response efforts and (1) catalyze investment and innovation in high-impact, scalable AI use cases, (2) promote ethical, responsible, and trustworthy AI development and use, (3) democratize access to AI technology and resources and (4) expand workforce AI capacity and capabilities. Through partnerships with stakeholders across the public health ecosystem, HHS can work toward a future where cutting-edge technologies such as GenAI-enabled chatbots to share basic health information and precision public health through deep-learning genomic algorithms help all Americans attain their highest level of health. HHS is committed to evolving its AI strategy as technologies and use cases continuously change in order to best improve the public's health.

[755] https://www.nimhd.nih.gov/programs/extramural/community-based-participatory.html

[756] https://www.cdcfoundation.org/community-based-organizations

6 Cybersecurity and Critical Infrastructure Protection

6.1 Introduction and Context

Securing digital systems from cyber threats is crucial for realizing the benefits and minimizing the risks of emerging technologies like AI. Without effective risk management, AI systems could put patient, participant, and public safety at risk, expose PII, and erode public trust in healthcare and public health systems. However, with appropriate controls, the possible benefits of AI to the nation's health and human services ecosystems are immense. Furthermore, addressing cybersecurity risks is essential to comply with E.O. 14410: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, which calls on HHS and the federal government to promote the safe and secure design, development, and deployment of AI models across critical infrastructure sectors. In response to the executive order and the National Cybersecurity Strategy,[757] HHS released its Cybersecurity Strategy[758] in December 2023, outlining actions to improve cybersecurity in health and human services. This document builds on HHS's Cybersecurity Strategy to highlight new actions the Department has taken since the Strategy's release and outline additional priorities.

The threat of cyber-incidents on the U.S. healthcare system is real and growing. Healthcare accounts for $4.5T (17%) of U.S. GDP and 9% of U.S. employment.[759] These factors contribute to making healthcare a large target. According to one survey, 92% of healthcare organizations experienced at least one cyber-incident in the past 12 months[760], and the HHS OCR reported a 264% increase in large data breaches involving ransomware from 2018 to 2022.[761]

In health and human services, cybersecurity incidents can impact multiple stakeholders. Previous incidents have led to delays in patient care and operational and financial disruptions for providers,[762] payers,[763, 764] and state public health departments.[765] Furthermore, the introduction of AI widens the threat landscape, as AI applications are increasingly used as tools for cyber attackers, exploitable vulnerabilities in digital systems, but also as new defensive tools. As AI adoption scales across the healthcare and public health ecosystem, cybersecurity protections must scale with it.

In this chapter, HHS outlines the current and expected trends in cybersecurity risks, how AI is impacting and creating these risks, and their implications for healthcare, public health, and human services. The Department then outlines the opportunities for actions to better enable organizations to address these threats, ongoing actions HHS has taken, and additional actions that could further bolster the health and human services ecosystem's cybersecurity capabilities.

---

757 https://www.whitehouse.gov/oncd/national-cybersecurity-strategy/

758 https://www.hhs.gov/about/news/2023/12/06/hhs-announces-next-steps-ongoing-work-enhance-cybersecurity-health-care-public-health-sectors.html

759 https://www.cms.gov/newsroom/fact-sheets/national-health-expenditures-2022-highlights https://www.bls.gov/spotlight/2023/healthcare-occupations-in-2022/

760 https://www.hipaajournal.com/92pc-us-healthcare-organizations-cyberattack-past-year/

761 https://www.hhs.gov/about/news/2024/09/26/hhs-office-civil-rights-settles-ransomware-cybersecurity-investigation-under-hipaa-security-rule-250-000.html

762 https://www.ucsf.edu/news/2020/06/417911/update-it-security-incident-ucsf

763 https://www.hipaajournal.com/change-healthcare-responding-to-cyberattack/

764 https://pmc.ncbi.nlm.nih.gov/articles/PMC7349636/ HIPAA journal is a US-based journal that provides comprehensive coverage of data breaches, guidelines for HIPAA compliance, and practical guidelines for data breach avoidance.

765 https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf

6.1.1 Action Plan Summary

Later in this chapter, HHS articulates proposed actions improve the sector's ability to manage its cybersecurity requirements. Below are the broad themes of these actions. For full details of proposed actions please see section 6.4 Action Plan.

Themes of actions:

1. Addressing the shortage of appropriately skilled cybersecurity workers to fill roles in health and human services
2. Supporting the standardization and alignment on best practices, especially in cybersecurity governance
3. Reducing and managing complexity in implementing new cybersecurity capabilities
4. Clarifying approach to navigate acute tensions between privacy and fairness and privacy and safety in health

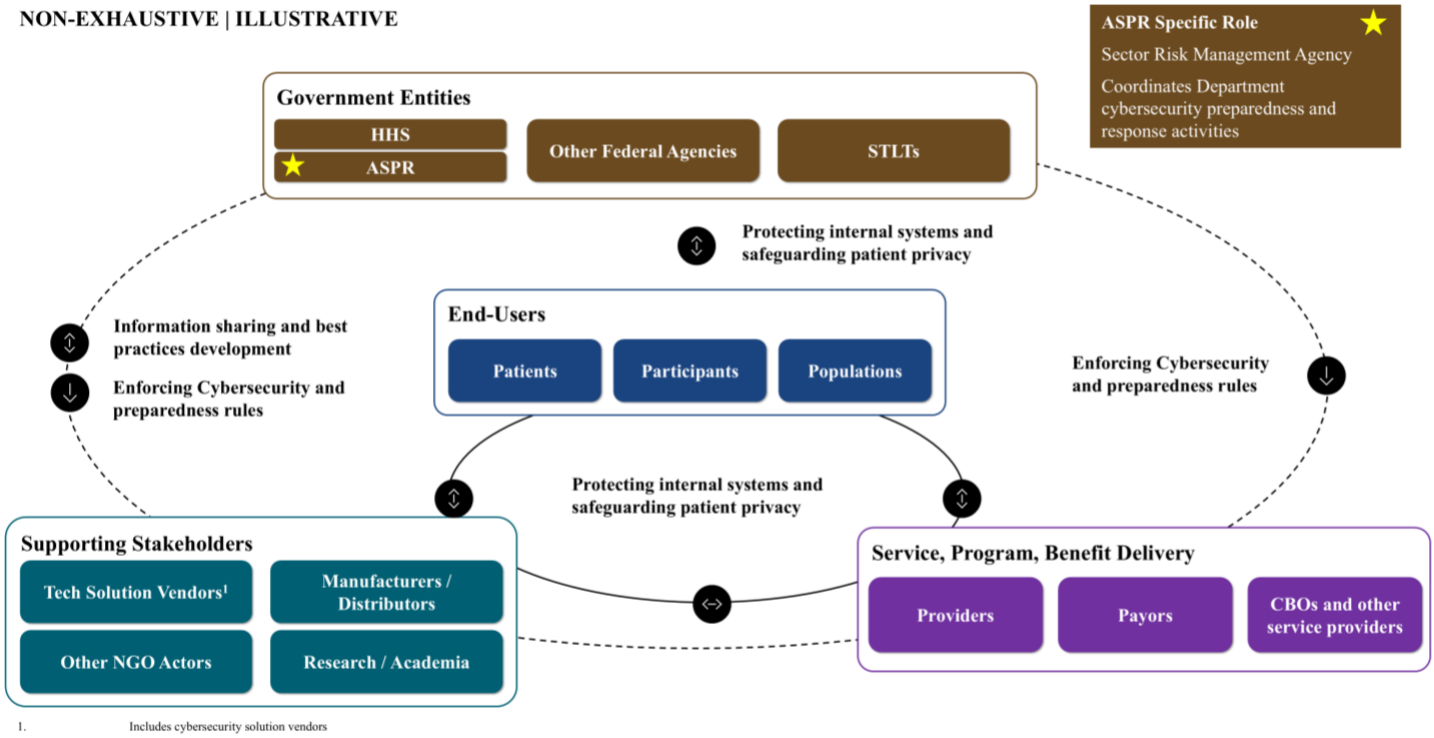
6.2 Stakeholders Engaged in the Cybersecurity and Critical Infrastructure in the Health and Human Services Ecosystem

HHS plays a dual role in promoting cybersecurity: first, by serving as a partner to the sector through information sharing and best practice development, and second, as a regulator, enforcing cybersecurity and preparedness rules. Alongside HHS, the rest of the federal government, STLTs, providers, payers, community organizations, and other non-government stakeholders are responsible for defending against cyber-threats and maintaining their organization's cybersecurity capabilities.

Multiple divisions and groups within HHS play a part in cybersecurity. These include the Health Sector Coordinating Council (HSCC) and HHS 405(d) Task Force, two public-private partnerships that aid in developing and sharing AI guidelines to healthcare, public health, and human services sector. ASPR coordinates Sector Risk Management Agency activities on behalf of HHS for the Healthcare and Public Health sector, coordinating cybersecurity preparedness and response activities within HHS, across the federal agencies, and with industry partners.

Exhibit 15 shows a non-exhaustive, illustrative diagram of example flows between stakeholders involved cybersecurity and critical infrastructure protection. Please note that the diagram does not capture all stakeholder roles and interactions. Please refer to other HHS documents for additional details on regulatory guidance and authorities. Roles may vary depending on domain or part of healthcare, public health, or human services ecosystem.

Exhibit 15: Interaction of Stakeholders in the Cybersecurity and Critical Infrastructure Protection Healthcare, Public Health, and Human Services Ecosystem.



6.3 Trends in Cybersecurity and Critical Infrastructure Protection

1. **Cyber-incidents are on the rise in the healthcare industry and globally, and the costs related to cybercrime are growing:** As companies, agencies, and organizations transform and modernize, the number and types of cyber threats grow each year. One analysis estimates a 589% increase in security vulnerabilities from 2023 to 2024 across industries.[766] Furthermore, the cost of cyber-incidents reached an estimated $8T in 2023 and continues to grow.[767] Health and human services organizations are also facing increased cybersecurity threats, including ransomware, phishing attacks, third-party breaches, data breaches, and social engineering attacks.[768] Accelerating digitization in healthcare (e.g., EHRs) has made healthcare a high-priority target for cyber-threats and magnified the complexity of establishing effective defensive measures. The health and public health sector saw a 42% increase in ransomware incidents between 2021 and 2022, and the frequency of cyber-incidents affecting health systems has doubled since 2016.[769] These incidents can cause system outages and endanger patient safety, among other consequences.
2. **The cybercrime industry is large and mature, with the capability to launch increasingly sophisticated attacks against health and human services organizations:** Cybercrime is a multi-billion-dollar, sophisticated industry replete with R&D functions that continuously improve their capabilities. Attackers are using new tools, including AI, to expedite the end-to-end attack life cycle from weeks to days or even hours. In recent years, attackers have used public health crises to demonstrate the power of their arsenal. For instance, during the COVID-19 pandemic ransomware and phishing attacks spiked globally due to a

766 https://info.jupiterone.com/scar-2023
767 https://www.usaid.gov/digital-development/cybersecurity/economic-growth-briefer
768 https://www.aha.org/h-isac-white-reports/2024-02-21-h-isac-tlp-white-announcement-h-isac-aha-executive-summary-cisos-current-and-emerging
769 https://aspr.hhs.gov/cyber/Pages/default.aspx

combination of increased threat activity and increased vulnerability due to the shift to work-from-home models.[770, 771]

3. **The use of AI, particularly GenAI, is expected to increase the number of cyber threats, vulnerabilities, and potential for errors and accidents:** The Federal Bureau of Investigation has warned that cybercriminals are increasingly leveraging AI tools with greater frequency to orchestrate targeted phishing campaigns.[772] For instance, AI-driven social engineering attacks, where AI impersonates a human using LLMs, are becoming increasingly successful.[773] In the first two months of 2023 alone, novel phishing attacks spiked 135%. Additionally, new malware is emerging that can evade traditional cybersecurity tools like endpoint detection and response (EDR) technology.[774, 775] For healthcare organizations, AI-driven phishing attacks are among the most used attack vectors in U.S. healthcare cyber threats.[776] In public health, AI-generated deepfakes can be used to spread misinformation, which can reduce people's willingness to seek treatment or simply undermine trust in public health institutions.[777, 778] Furthermore, AI-powered systems could also be used to de-anonymize sensitive health information, leading to costly ransomware attacks.[779] This is particularly troubling given the wide range of anonymized datasets available in healthcare and public health for clinical trials, precision medicine, and medical research. Other healthcare specific threats could include vectors like adversarial attacks on medical imaging[780] or data poisoning,[781] while threats affecting federal agencies or STLTs include automated social engineering attacks[782] or disinformation campaigns. A broad range of other adversarial AI techniques exist in various stages of development and sophistication.[783]
4. **The increasing need for and access to large datasets in health and human services is also leading to a greater risk of data breaches:** While healthcare lags other sectors in adopting cloud storage, the increase in online patient platforms, AI adoption, and EHR use led to more health data being stored in the cloud.[784] Healthcare, public health, and human services organizations manage large, sensitive datasets, including PHI, and many stakeholders have access. As data increasingly migrates to cloud storage, all organizations must take cybersecurity precautions to safeguard sensitive data. Furthermore, relying on third-party cloud storage can magnify vulnerabilities. In fact, 35% of healthcare data breaches involve third-party vendors.[785] The ramifications of data breaches in healthcare are immense. For example, one ransomware attack in February exposed the private health information of 100 million individuals and may have resulted in a financial impact exceeding $2.5B.[786, 787] Healthcare data breaches are increasingly costly due to losses from business

[770] https://pmc.ncbi.nlm.nih.gov/articles/PMC9212240/

[771] https://pmc.ncbi.nlm.nih.gov/articles/PMC9755115/

[772] https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence

[773] https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence

[774] https://www.hhs.gov/sites/default/files/ai-cybersecurity-health-sector-tlpclear.pdf

[775] https://www.hyas.com/blog/blackmamba-using-ai-to-generate-polymorphic-malware. An example of this is the Black Mamba polymorphic malware which dynamically modifies its behavior to avoid detection.

[776] https://www.hipaajournal.com/healthcare-data-breaches-due-to-phishing/

[777] https://www.nyu.edu/life/information-technology/safe-computing/protect-against-cybercrime/ai-assisted-cyberattacks-and-scams.html

[778] https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf

[779] https://www.hipaajournal.com/managed-care-of-north-america-hacking-incident-impacts-8-9-million-individuals/

[780] https://pmc.ncbi.nlm.nih.gov/articles/PMC10487122/ Manipulate medical images in a way that deceives diagnostic systems, leading to misdiagnosis or incorrect treatment decisions.

[781] https://pmc.ncbi.nlm.nih.gov/articles/PMC10984073/ Attackers manipulate training data in an AI model by injecting false data, leading to biased models or inaccurate output.

[782] https://www.weforum.org/stories/2024/10/ai-agents-in-cybersecurity-the-augmented-risks-we-all-need-to-know-about/ Using personalized messages to convince someone to divulge sensitive information or click a malicious link.

[783] Other AI-enabled cyber-attacks include generating deceptive AI (e.g., deepfake attacks, morphing attacks), attacks on AI systems (e.g., data poisoning, evasion attacks, model extraction), emerging technologies (e.g., quantum computing, false biometric data), and dual-use AI capabilities (e.g., computer vision, NLP, audio recognition.

[784] https://www.hipaajournal.com/healthcare-cloud-usage-grows-but-protecting-phi-can-be-a-challenge/

[785] https://www.hipaajournal.com/healthcare-highest-third-party-breaches/

[786] https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf Incident logged on July 19, 2024.

[787] https://www.hipaajournal.com/change-healthcare-responding-to-cyberattack/

disruption, customer support, and remediation. The average cost of a healthcare data breach for an organization is now $10M.[788, 789]

5. **Traditional tools for combatting cyber-threats are still effective, but cyber risks are outpacing capabilities in organizations due to several challenges:** Although AI-enabled cyber threats can be more devastating, they often have vectors that resemble traditional cybersecurity attacks. Recent data shows that 84% of critical infrastructure incidents involve, "an initial access vector that could have been mitigated with best practices and security fundamentals, such as asset and patch management, credential hardening, and the principle of least privilege."[790] Traditional cybersecurity practices can still help thwart these threats. However, organizations are struggling to implement even traditional tools for combatting cyber-threats due to challenges such as a mismatch of skillsets in cybersecurity workforce, lack of standardization of best practices, implementation complexity, and other barriers. In the next section, HHS provides additional context to those challenges and outlines opportunities for the Department to take action to enhance the sector's cybersecurity and critical infrastructure protection.

6.4 Action Plan

Health and human services organizations are investing more in their cybersecurity capabilities. One estimate suggests that the global healthcare industry will spend $125B on cyber products and services from 2020-2025, representing 15% annual growth.[791] and a 2023 survey of healthcare cybersecurity professionals found that over half had seen increases in their cybersecurity budgets in the past year.[792] Despite increased attention and investment, healthcare organizations are struggling to keep up with escalating threats. For instance, ransomware attacks on the healthcare sector nearly doubled from 2022 to 2023.[793] Moreover, the focus of cybersecurity spending has shifted; from 2016 to 2022, the share dedicated to preventing incidents decreased from 60% to 30%, with more resources now allocated to managing ongoing incidents.[794]

Below, HHS outlines several opportunities for actions to improve the sector's ability to manage its cybersecurity requirements. These opportunities are:

1. Addressing the shortage of appropriately skilled cybersecurity workers to fill roles in health and human services
2. Supporting the standardization and alignment on best practices, especially in cybersecurity governance
3. Reducing and managing complexity in implementing new cybersecurity capabilities
4. Clarifying approach to navigate acute tensions between privacy and fairness and privacy and safety in health and human services

For each of these opportunities, HHS has added context and highlighted where it has taken mitigating actions and where it is considering future action.

[788] https://aspr.hhs.gov/cyber/Pages/default.aspx

[789] https://www.ibm.com/reports/data-breach Global average cost for a data breach is $4.88 million, for comparison.

[790] https://www.ibm.com/downloads/documents/us-en/107a02e952c8fe80

[791] https://www.hipaajournal.com/healthcare-cybersecurity/

[792] https://www.chiefhealthcareexecutive.com/view/healthcare-cybersecurity-budgets-are-rising-but-workers-are-hard-to-find.

[793] https://www.dni.gov/files/CTIIC/documents/products/Ransomware_Attacks_Surge_in_2023.pdf

[794] https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/cybersecurity/cybersecurity-trends-looking-over-the-horizon

1. **Addressing the shortage of appropriately skilled cybersecurity workers to fill roles in health and human services.**

Context:

Many organizations lack the cybersecurity talent, knowledge, and expertise required to defend against the latest threats and are struggling to fill essential roles. Across the U.S., the gap in skilled cybersecurity workers is widening faster than new hiring can keep up.[795, 796] This shortage may stem from a mismatch in skillsets rather than a lack of job-seeking cyber professionals. More traditional cyber professionals do not have the required expertise in areas like cloud services, AI and GenAI data and analytics, or health and human services IT. Increasingly, leaders outside of cybersecurity teams are also searching for cybersecurity talent. However, leaders often lack the basic understanding of cybersecurity needed to evaluate candidates or meet their hiring needs effectively. In the healthcare sector, one survey by CDW revealed that only 14% of healthcare IT leaders reported having fully staffed security teams.[797] HHS has taken actions to improve cybersecurity workforce capabilities and, as outlined below, will look to develop trainings and explore resourcing to bring more appropriately skilled talent into the sector.

HHS actions to date (non-exhaustive):

- **Increasing capabilities for under-resourced STLTs through active monitoring, data sharing, and collaboration.** HHS continues to monitor and share data, including for AI threats, to increase the capabilities of under-resourced STLTs and work with its government partners to develop and share draft guidelines on essential cybersecurity practices to protect AI models and continue providing tools and resources to help under-resourced entities implement robust cybersecurity practices.
- **ARPA-H is developing new tools that automatically detect and fix cyber vulnerabilities**, reducing the cybersecurity burden on hospitals and healthcare organizations. These steps include:
    - **Launching AI Cyber Challenge** in collaboration with DARPA to leverage AI to create usable, automatic tools for vulnerability identification and remediation that can be deployed across the Nation's open-source software supply chain.
    - **Creating Universal Patching and Remediation for Autonomous Defense program,** which intends to develop an autonomous cyber-threat solution that enables proactive, scalable, and synchronized security updates, reducing the uncertainty and manual effort necessary to secure hospitals.

HHS near-term priorities:

- Develop additional health- and human-services-sector-specific cybersecurity training geared toward organizational leadership and hiring managers outside cyber teams.
- Assess opportunities to support cybersecurity workforce development for under-resourced healthcare and public health organizations.
- Support adoption of technologies in HHS, STLTs, and CBOs that support secure data sharing.

HHS long-term priorities:

- Integrate new cybersecurity requirements in HHS grants, contracts, and cooperative agreements.
- Explore incorporating AI-enabled threats into HHS Priority Intelligence Requirements to increase existing sharing of cyber threat intelligence across HHS and healthcare, public health, and human services sectors.

[795] https://www.weforum.org/stories/2024/04/cybersecurity-industry-talent-shortage-new-report/

[796] https://www.whitehouse.gov/oncd/briefing-room/2024/09/04/service-for-america-cyber-is-serving-your-country/

[797] https://www.hipaajournal.com/healthcare-cybersecurity/

2. Supporting the standardization and alignment on best practices, especially in cybersecurity governance

Context:

Cybersecurity comprises a complex set of capabilities from strategy to data protection to resilience and recovery. Each organization values and prioritizes its cyber capabilities differently. As a result, there are no accepted standards for when and how to use trusted architecture techniques. In the health sector, these challenges could extend to cyber-related risk governance, where, sometimes, there are poorly defined roles and responsibilities for addressing failures in AI systems, a lack of understanding of liability when AI systems are used for decision-making, and an inability to validate model outputs.[798] In many organizations, this can lead to an ad hoc approach to cyber management. Often the responsible cybersecurity team, normally the Chief Information Security Officer (CISO) for IT-related threats, is not consistently provided with the resources they need to protect their organization effectively. Additionally, they might not be involved early enough to assess cyber risks for new programs or may only be brought in when a breach occurs. Furthermore, this can lead to a gap in cyber capabilities, including asset management, vulnerability management, impactful metrics and reporting, identity and access, and data protection. HHS has taken action to adopt and publish best-practice standards and will continue to develop and update guidelines as cybersecurity practices evolve.

HHS actions to date (non-exhaustive):

- **Adopted the NIST AI Risk Management Framework:**[799] integrated AI-risk-management practices into planning for cybersecurity, emergency management, clinical operations, medical devices, legal, workforce management, supply chain, and procurement.[800]
- **Released HHS's Cybersecurity Strategy (December 2023)**[801] recommended implementing basic, traditional cybersecurity measures and is geared toward helping all organizations elevate their security floor with the proper tools and measures to manage the risks of AI in their organization.
- **Released its HPH Cybersecurity Performance Goals (CPGs)**[802] help organizations prioritize the implementation of high-impact cybersecurity practices and are adapted from CISA's own CPGs[803] and from best practices in the industry to fit the healthcare context.
- **Released proposed measures to strengthen cybersecurity in healthcare under HIPAA (December 2024)** by requiring health plans, healthcare clearing houses, and most health providers and their business associates to better protect individuals' electronic PHI against both external and internal threats.
- **Collaborates with government partners** to develop and share draft guidelines on essential cybersecurity practices to protect AI models.

HHS near-term priorities

- Develop guidelines on maintaining operations after a system deploying AI is compromised. Users should have policies, tools, and training in place to understand when AI systems are producing incorrect outputs, and resiliency plans for when AI systems are compromised.
- Update existing regulations and guidelines on adoption to include best practices for maintaining cybersecurity, including for maintaining secure means of data transfer and sharing

[798] https://healthsectorcouncil.org/health-industry-cybersecurity-artificial-intelligence-machine-learning/

[799] https://www.nist.gov/system/files/documents/2022/08/18/AI_RMF_2nd_draft.pdf

[800] https://www.hhs.gov/sites/default/files/public-benefits-and-ai.pdf

[801] https://www.hhs.gov/about/news/2023/12/06/hhs-announces-next-steps-ongoing-work-enhance-cybersecurity-health-care-public-health-sectors.html

[802] https://hhscyber.hhs.gov/performance-goals.html

[803] https://www.cisa.gov/cybersecurity-performance-goals

3. Reducing and managing complexity in implementing for new cybersecurity capabilities

Context:

Threats described above point to the need for stronger data security and access controls at each stage of AI application development. Secure design and training can help prevent AI confabulation, data breaches, and data exfiltration, which can be particularly important for institutions conducting sensitive research on biotechnology, new pathogens, and more. Developers of AI models must implement robust security controls into each facet of their operations, and users of those solutions (e.g., healthcare stakeholders) need training to understand and safely integrate those solutions.[804] However, given persistent staffing shortages, a lack of standardization, and an increasingly complex technology landscape, organizations struggle to coordinate with solution providers, shift their organizational norms to comply with new deployments or deploy and scale solutions across the entirety of their enterprises. The Department will identify ways to reduce the complexity through potential actions outlined below such as enhancements to HHS healthcare IT certifications.

HHS actions to date (non-exhaustive):

- **Issued guidelines to enhance software transparency.** Section 524B(b)(3) of the FD&C Act requires that medical device manufacturers of "cyber devices" provide a software transparency mechanism called a "Software Bill of Materials" (SBOM) as part of their premarket submissions. The SBOMs serve as one part of FDA's evaluation of device security, postmarket vulnerability, and incident response. FDA will continue to monitor device AI cybersecurity considerations in premarket submissions and postmarket issues to assess whether additional policy is necessary to safeguard patient safety.
- **The Digital Health Security (DIGIHEALS) Program is working with AI and cybersecurity experts to strengthen our electronic health ecosystem** by adapting proven technologies developed for national security so those technologies can be used in civilian health systems, clinical care facilities, and even personal health devices.

HHS near-term priorities:

- Encourage Health IT developers to implement privacy and security by design in their products, including building cyber controls into products or offering service APIs to integrate cyber controls into other systems. This can be achieved by:
    - Enhancing cybersecurity certification criteria in ASTP's ONC Health IT Certification Program.
    - Broadening ASTP's ONC Health IT Certification Program's scope to include additional health IT systems (e.g., laboratory systems, telemedicine—patient health records, exchange, or access systems, clinician-led clinical data registries, electronic prior authorization systems, and clearinghouse processing systems).
- Partner with healthcare stakeholders to develop guidelines and resources for organizations looking to assess the risks of AI to their organization. This can include acquisitions and procurement guidelines to help small/under-resourced organizations assess the security impact of different AI tools and solutions, train staff on best practices, or for assessing risk for data transfer and data-sharing tools that are considered secure.
- Map security risks across HHS value chains of AI systems for security and privacy risks to help address third-party concerns, including corrupt libraries, unvetted data, or label errors.[805]
- Provide funding to research the impact of cybersecurity on clinical settings.

HHS long-term priorities:

- Consider partnership with industry and other government agencies to develop Key Risk Indicators and performance thresholds that enhance software transparency.

[804] https://www.rand.org/pubs/research_reports/RRA2849-1.html

[805] https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

4. **Clarifying approach to navigate acute tensions between privacy and fairness and privacy and safety in health and human services.**

   Context:

   Cybersecurity teams in health and human services contexts need to balance the desire for privacy and data protection with broader goals that are sometimes in contradiction. For instance, an organization may want to enforce strict data privacy and data-sharing restrictions while also aiming for broad data inclusion in AI models to mitigate potential bias or monitor population health. In practice, without standards or guidelines, these conflicting priorities can lead to delayed implementation decisions or sub-optimal design choices that neither sufficiently protect privacy nor lead to broader health-related goals. HHS will work to try to clarify and frame cybersecurity trade-offs to assist stakeholders in making decisions.

   HHS near-term priorities:

   - Provide guidelines on how organizations can navigate questions of cybersecurity trade-offs and where to focus most on protecting cybersecurity and privacy.

6.5 Conclusion

In healthcare, public health, and human services, disruptions from cyber threats directly impact lives. It is crucial for HHS and its broader ecosystem to recognize the growth of cyber threats and take steps to ensure their organizations are protected. Cybersecurity is a fundamental capability for any organization in the broader HHS ecosystem that is looking to expand its use of AI applications responsibly and effectively. However, as noted above, despite widespread awareness of the threat and increased focus on cybersecurity, organizations are struggling to keep pace with potential attackers. There remains significant opportunity to improve skillsets of cybersecurity talent, establish and promote standards for best practices and governance, reduce the complexity of implementing new capabilities, and assist organizations in balancing questions of cybersecurity and privacy.

HHS's Cybersecurity and Infrastructure Protection strategy will continue to evolve as the threat landscape changes. HHS is committed to assisting its agencies and the broader HHS stakeholders in improving their cybersecurity capabilities and has taken several steps to do so. The Department will continue to consider additional actions to lift the security floor for the healthcare system and to make it easier and safer for organizations to adopt AI applications that positively impact the American people.

7 Internal Operations

7.1 Introduction and Context

AI presents wide-ranging opportunities for the Department. HHS operating and staff divisions have already been using AI to advance their missions to improve internal operations and enhance the execution of public-facing services. The scale at which AI is used across HHS requires a formal, departmentwide approach. The Department's approach to AI must also focus on change management and adaptability, as AI implementation and use can transform existing processes. By optimizing Department processes, policies, and structures for procuring, testing, deploying, and securely managing AI solutions internally, HHS aims to accelerate knowledge sharing and coordinate support of AI investments. This will ensure greater consistency across the Department while also allowing for appropriate agency-level flexibility to drive innovation.

In alignment with E.O. 13859, E.O. 14110, and OMB Memoranda M-24-10 and M-24-18, HHS's Office of the Chief Artificial Intelligence Officer (OCAIO) will lead three focal areas needed to deploy high-value, trustworthy AI within HHS, both at the Department level and within HHS's divisions:

1. **Governance**
2. **Internal process improvement and innovation**
3. **Workforce and talent**

To create a cohesive strategy, these focal areas must be integrated into the major internal operations of HHS divisions and implemented at all departmental levels. This approach will help appropriately balance centralized coordination from the OCAIO with the necessary flexibility needed by HHS divisions to achieve their respective mission goals. Additionally, the Department's AI Strategy will align with existing policies, frameworks, and statutory responsibilities for IT infrastructure review and deployment.

7.2 Opportunities and Risks

Opportunities:

Increasing AI adoption and use within the Department's internal operations presents significant opportunities. These include:

1. **Improving quality, experience, and safety of public-facing programs and services:** AI can enable HHS to more effectively deliver health and human services to hundreds of millions of individuals each year by deploying use cases that have been appropriately validated in the private or public sector.[806] For example, HHS agencies involved in the direct provision of patient care can leverage technologies for more accurate patient monitoring, and agencies involved in the delivery of human services can use AI to connect beneficiaries with best-fit services in a more efficient way.[807]
2. **Informing policy, guidelines, and processes that support innovation and safe use of AI *within HHS:*** HHS will need to keep pace with a rapidly evolving technology ecosystem to successfully execute its mission. Piloting and deploying use cases that assist in setting effective guidelines and improving processes will enable HHS to operate more effectively, which in turn enables the Department to best provide oversight and delivery of health and human services in the U.S.
3. **Building knowledge and capabilities to inform public-facing policy and guidelines for HHS domains:** Internal adoption of AI for HHS operations will increase the department's AI knowledge and capabilities. This, in turn, allows HHS to provide more informed policy, guidelines, and oversight of these technologies in HHS's domains (e.g., healthcare delivery and R&D).
4. **Improving workforce efficiencies:** AI has the potential to automate current manual processes that require direct human staff and contractor time. Leveraging AI to facilitate tasks that can be augmented through technology will allow HHS staff to spend more time performing high-impact activities (e.g., review, coordination, enforcement, and direct provision of care where applicable).

Risks:

The use of AI to support HHS's mission increases some existing risks while introducing new risks. Examples of such internal risks include:

- **Data privacy and security:** As HHS divisions gain experience with AI solutions, it is possible that future use cases may use sensitive patient-, participant-, or community-level information to train internal models or produce outputs. These types of uses will need to be managed with the same vigilance as non-AI use cases and may potentially require additional controls or adoption of other technologies depending on the context and data source to ensure that AI use is not creating new vulnerabilities.
- **Execution risk:** Overly strict internal guidelines on exactly when AI may or may not be used risks disincentivizing innovative approaches that could create a positive impact. Research demonstrates that preconceptions about AI and its impact (e.g., on careers or on program participants) can hinder the successful deployment of new AI technologies in the workplace.[808, 809] Poor communication with staff at HHS about AI progress and challenges may further exacerbate this risk.
- **Impact on workforce training and skills:** Integrating AI into the technical and workforce workflows of HHS divisions opens the door to risks, such as a skills gap if individuals no longer perform tasks that were once part of their scope. This may not be a risk for rote tasks (e.g., calculation, basic arithmetic, scheduling) but may pose challenges for skills like customer support or other human interaction. Additionally, the use of AI may lead to overreliance, where the staff responsible for overseeing the functions supported by AI fail to exercise sufficient oversight.

These risks will be considered as part of the proposed actions in Planned HHS Activities below. Successful execution of this Strategic Plan faces additional risks if not appropriately managed. In addressing these or other risks posed by the development and or use of AI, HHS will additionally tailor risk management strategies to the anticipated level of risk associated with a specific model, tool, or use case. Systems incorporating AI should apply

[806] https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/doc21-ed-508.pdf

[807] https://www.nejm.org/doi/full/10.1056/NEJMra2204673

[808] https://pubmed.ncbi.nlm.nih.gov/37927664/

[809] https://english.rekenkamer.nl/publications/reports/2024/10/16/focus-on-ai-in-central-government An international example from Netherlands Audit on use of AI in government finding comprehensive AI assessments created significant cost or time requirements disincentivizing use and deployment.

risk management practices to identify, address, and monitor potentially negative impacts through all phases of relevant processes and system frameworks.

7.3 Governance

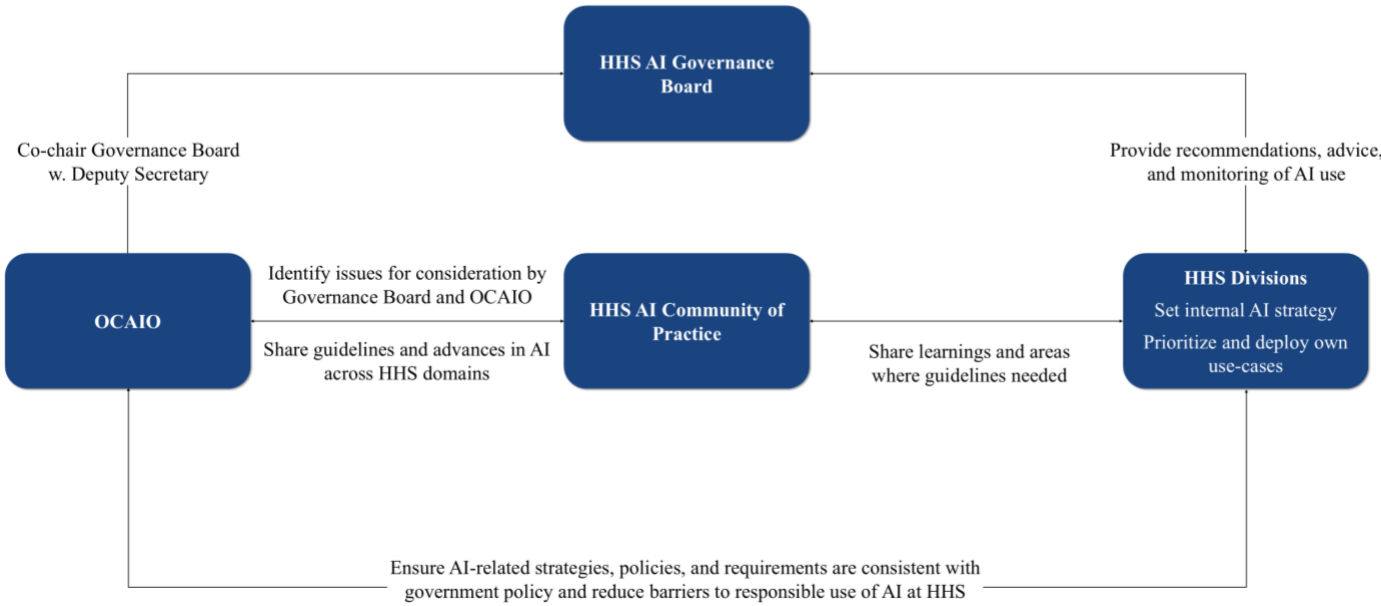
Context:

Effective AI governance throughout the entire solution life cycle—from conceptualization of an AI intervention to execution and decommissioning of the tool—is essential to facilitate appropriate adoption and risk management. Just as HHS and its divisions have built significant infrastructure around IT to ensure responsible use and minimize risks from improper data sharing, HHS has and will continue to put into place necessary safeguards around AI. AI governance practices will leverage existing HHS and/or division-level governing bodies and processes where possible to ensure strategic alignment and avoid undue burden on HHS divisions. AI governance will take a tailored approach to each division's unique structures and needs to promote innovation while minimizing the potential impacts of AI-related risks.

Exhibit 16 shows the interaction of the OCAIO and HHS agencies and defines at a high level the OCAIO role within HHS. Governance mechanisms illustrated in the Exhibit are further detailed below.

Exhibit 16: Interaction of OCAIO and HHS agencies

NON-EXHAUSTIVE | ILLUSTRATIVE



HHS actions to date (non-exhaustive):

HHS has already created foundational governance structures to support the use of AI, including:

- **Hired a permanent CAIO** consistent with M-24-10's requirements. The OCAIO will ensure that all strategic and Department policies, requirements, and guidelines are consistent with government policy and reduce barriers to the responsible use of AI. While the OCAIO holds primary responsibility for the governance of AI solutions across HHS, the OCAIO will consult and collaborate with other offices in the Department to ensure their broad applicability to HHS divisions.
- **Created the HHS AI Governance Board** to serve as the principal governance body responsible for guiding HHS's AI policies, programs, and technology uses and ensuring that these policies are aligned to FAVES principles. It provides recommendations, advice, and monitoring on key issues surrounding AI

use. The Board first met in May 2024, is chaired by the Deputy Secretary, co-chaired by CAIO, and is comprised of senior leaders from HHS divisions. It is responsible for supporting AI governance, developing strategic AI priorities across the enterprise, and overseeing strategic execution. The Board will also monitor progress toward HHS's implementation of this Strategic Plan.

- **Created the HHS AI Community of Practice (CoP)** run by the OCAIO that includes AI-interested staff from across HHS. The goals of the CoP are to provide an opportunity for ongoing learning and collaboration across the Department, surface priority issues for HHS-level and cross-agency coordination by the OCAIO and help identify key issues for consideration by the HHS AI Governance Board. The HHS AI CoP also supports workgroups in key topic areas like AI policymaking and AI talent and workforce development.

HHS near-term priorities:

While HHS has developed the governance approach outlined above, it will take time to refine and implement the more comprehensive structures and processes needed to accelerate the adoption of use cases and ensure organizational readiness for AI-driven mission enhancements. To further develop these governance practices, the HHS OCAIO will:

- **Strengthen and formalize a comprehensive governance structure:** HHS will expand upon the foundational elements detailed above to support the responsible use of AI within the organization. This will include articulating and documenting roles, responsibilities, and decision rights across key governance bodies.
- **Provide guidelines to HHS divisions on governance to implement AI within their scope:** These guidelines will support the development of any division-specific governance policies needed to execute each agency's unique missions and will consider the existing structures they have established.
- **Enrich the Community of Practice by stewarding AI working groups:** These groups will share real-time insights on the use of AI across the Department, including shared learning, best practices, and additional avenues for confirming emerging issues. HHS will additionally continue to explore ways to expand the CoP over time to suit the Department's needs.
- **Establish a regular cadence for reviewing and revising AI governance structures:** The HHS AI Governance Board will establish a process for regularly reviewing HHS AI structures and guidelines (e.g., annually).

7.4 Internal Process Improvement and Innovation

Context:

HHS must ensure that its processes are set up in a way that facilitates the safe and effective use and development of AI. This spans multiple types of workflows, including acquisitions and procurement (e.g., procurement of AI solutions, use of AI in selecting bespoke tools), prototyping, piloting, and deployment (e.g., creation of analytics engines for disease prevalence monitoring), maintenance and operations (e.g., ensuring ongoing quality and compliance), and security (e.g., avoiding misuse of sensitive data). Similarly, HHS must align its internal processes for grant-making, grant oversight, and program evaluation as needed to align to best practices for adoption of AI where applicable and maintain programmatic and scientific integrity and sustainability. HHS already has multiple policies guiding each of these areas, and AI uses will need to remain aligned with existing approaches. For example, sensitive data storage must still be held to the same high bar whether AI is used or not.[810, 811] In developing its approach, HHS will evaluate whether to update existing policies (e.g., Authority to

[810] https://uscode.house.gov/view.xhtml?path=/prelim@title44/chapter35/subchapter2&edition=prelim 44 U.S.C. §§ 3551 et seq (FISMA)

[811] https://www.whitehouse.gov/omb/management/ofcio/m-24-15-modernizing-the-federal-risk-and-authorization-management-program-fedramp/ OMB M-24-04, OMB M-24-15

Operate frameworks) to ensure they support the use and development of AI. These policies will also follow OMB M-24-10 and M-24-18 and will include relevant additional steps to identify and mitigate risks, such as when an AI solution has been deemed "rights impacting" or "safety impacting."

HHS actions to date (non-exhaustive):

HHS has already set the foundation for the use of AI, including:

- **Compiled the AI Use Case Inventory,** in accordance with EO 13960, and provided a public inventory of non-classified and non-sensitive current and planned AI use cases. This inventory details ways in which HHS can leverage AI and includes oversight methodologies and benefits. In 2024, the AI Use Case Inventory included 271 use cases across 13 agencies. EO 13960 initiated this use case library, which EO 141110 later endorsed and enhanced. HHS will update the inventory annually, consistent with the new requirements expressed in OMB Memo M-24-10.

HHS near-term priorities:

In addition to cataloging the Department's AI use cases, HHS intends to build the necessary internal processes and support structures to enable the adoption of responsible AI. To this end, the HHS OCAIO will:

- **Coordinate the development of enterprise AI procurement approaches and toolkits:** This work will provide guidelines at the HHS level to promote a standardized approach to procuring AI tools, technologies, and subject matter expertise and will be designed in close collaboration with HHS divisions to ensure it provides sufficient guidelines across HHS and remains aligned to Federal Information Technology Acquisition Reform Act requirements.[812] The HHS OCAIO will additionally explore the inclusion of AI-specific language into the HHS Acquisition Regulations[813] and other relevant policies.[814]
- **Support responsible prototyping and piloting:** This support will include establishing, co-leading, and funding pilots at both the Department and division levels to address enterprise solutions applicable to numerous HHS divisions and unique mission-specific uses. The HHS OCAIO intends to help facilitate the establishment and use of "AI sandboxes" for rapid prototyping and solution evaluation (e.g., testing whether an algorithm leads to the desired result), and security assessments (e.g., will deployment of algorithm exacerbate or create new cybersecurity risks for HHS) prior to cross-department deployment.
- **Ensure oversight for AI quality monitoring:** Ultimately, HHS and divisions will be responsible for ensuring the compliance of AI with applicable standards. The HHS OCAIO will implement monitoring systems for AI solutions at the department level (including accuracy, reliability, and traceability) and will advance and support capabilities for monitoring AI tools. The OCAIO will additionally issue, as applicable, guidelines to HHS divisions for establishing division-specific monitoring systems for their agency use.
- **Ensure oversight and update processes to promote AI security:** Distinct from the quality monitoring above, the HHS OCAIO will work with the Office of the Chief Information Officer (OCIO) to ensure AI use cases meet applicable security requirements. Consistent with its responsibilities, OCIO will follow its processes to ensure that IT utilizing AI is properly secured and will update security processes as needed to reflect the changing AI landscape. The OCAIO and OCIO will collaborate with other HHS stakeholders to ensure these processes can be applied across HHS divisions.
- **Issue guidelines on use of AI:** The HHS OCAIO will provide guidelines to help HHS divisions determine when and under what circumstances it makes the most sense to use AI solutions. The

[812] https://www.cio.gov/handbook/it-laws/fitara-2014/

[813] https://www.hhs.gov/grants-contracts/contracts/contract-policies-regulations/hhsar/index.html

[814] https://www.hhs.gov/grants-contracts/contracts/contract-policies-regulations/hhsar/index.html

guidelines will also include the specific steps that must be taken for rights—and safety-impacting AI use cases and other AI-use cases as needed[815] consistent with EOs, OMB M-24-10, and other guidelines.

7.5 Workforce and Talent

Context:

The goal of an AI-enabled workforce is to allow individuals to perform their duties safely and effectively, leveraging AI tools where reasonable to assist in their workflows. HHS will continue to evaluate opportunities to leverage AI in daily workflows and aims to be responsive to a dynamically changing technological landscape in the Department. In certain scenarios, AI can optimally allow individuals to reallocate their time to the highest-impact areas, for example, by minimizing time spent on manual data analysis and spending more time on decision-making and program improvement.

Existing federal actions:

Other federal agencies have already prioritized enabling workforce and talent using AI, namely by:

- **Developed the Office of Personnel Management's (OPM) "Workforce of the Future" playbook** in February 2024 which details workforce strategy and offers guidelines for federal agencies on the integration of AI. In particular, the playbook includes several calls to action for federal agencies, including leveraging appropriate AI capabilities into HR processes, understanding how AI will impact the workforce, upskilling teams with appropriate competencies, and training the workforce on AI use cases.[816]
- **Included AI roles within OPM's Direct Hire Authority (DHA) framework** in December 2023 which allowed federal agencies to bypass specific hiring processes for high-demand fields. This strategic decision enables agencies to attract and hire skilled AI specialists who can meet complex agency requirements without traditional hiring procedures that may otherwise deter them from joining government agencies. Building on this authority, HHS has developed standard AI Position Descriptions to increase hiring speed using the DHA across the Department.
- **Piloted GenAI to enable workforce to automate previously manual data analysis that informs decision-making and program improvement.** ASTP's CAIO and Office of Policy are exploring how GenAI can streamline the end-to-end process of managing, analyzing, and incorporating public comments during federal rulemaking. The focus is on using engineered prompts to produce usable comment summaries for the Office of Policy's rulemaking activities.

HHS near-term priorities:

To establish an AI-enabled workforce, the HHS OCAIO will:

- **Collaborate with governmentwide leaders to develop an AI hiring strategy:** The HHS OCAIO will collaborate with other government stakeholders (including the OPM and Office of Management and Budget) to develop a strategy for hiring skilled AI specialists. This strategy could include identifying AI needs at the Department level, evaluating pay scales for AI roles, and establishing shared resources to be used across federal entities (e.g., AI-related position descriptions).
- **Collaborate with HHS leaders to perform a gap assessment of AI skills:** The HHS OCAIO will collaborate with the HHS Chief Human Capital Officer and other division workforce leaders to perform a gap assessment of the Department's current workforce AI capabilities. This will identify areas for targeted intervention, which may include upskilling current talent or recruiting new talent (either within

[815] Controls should be in place even where AI is not rights/safety impacting, for example, individual data protection controls, IP rights controls, contractual compliance, records management, and protection of CUI/procurement sensitive/trade secrets/other non-individual data, among other considerations.

[816] https://www.opm.gov/workforce-of-the-future/wof-playbook.pdf

the federal government or externally) with these skills. The gap assessment's output will additionally inform a funding plan for closing identified gaps.

- **Improve AI literacy for all HHS staff:** In addition to the gap analysis, the HHS OCAIO will facilitate the delivery of foundational AI literacy training to help all HHS staff and contractors become more comfortable with AI and share an understanding of the potential benefits, limitations, and risks of AI technologies.

7.6 Conclusion

In this chapter, HHS outlined the steps the Department has taken and will continue to take in the future to realize the benefits of AI in its internal operations and stay nimble and current with the rapidly evolving AI landscape. HHS recognizes that the transformative potential for AI extends to its own internal operations and not just to the work of its divisions and of the many stakeholders of the health and human services ecosystem. HHS sees significant opportunity for AI to improve its public-facing programs and services, improve processes that support innovation at HHS, inform policy and guidelines, and improve workforce efficiencies. These opportunities, if responsibly undertaken, could enable the Department to better fulfill its mission of improving the health and well-being of the American people.

Conclusion

HHS aims to be a global leader in innovating and adopting responsible AI to achieve unparalleled advances in the health and well-being of all Americans. This Strategic Plan outlines the ways in which HHS intends to achieve that goal.

The use of AI in medical research and discovery, medical product development, safety, and effectiveness, healthcare delivery, human services delivery, public health, cybersecurity, and HHS's operations is no longer a speculative future but a present reality, driven by rapid technological advancements. In recent years, AI has become part of everyday life, including within the health, human services, and public health ecosystem. This evolution is evident in the ability of AI to serve as a tool that supports delivering high-quality care, streamlining drug development, speeding and improving health and human services communications, and more.[817] Moreover, AI can enhance health equity, for example through providing real-time, automated translation services for individuals facing language barriers or supporting individuals with disabilities through optimized speech patterns and fluent conversation.[818] The use of AI brings these and many additional promising benefits discussed throughout the chapters of this Strategic Plan, yet comes with a wide range of risks such as the potential for AI to propagate biases, misclassify patient needs, or breach confidentiality.

HHS is dedicated to not only fostering the adoption of AI to achieve enhanced outcomes but also protecting patients, caregivers, and all stakeholders from these and other potential pitfalls discussed in each chapter of the Strategic Plan. This commitment involves implementing robust measures to address these challenges while promoting the transformative potential of AI.

As AI continues to evolve rapidly, HHS will adopt an equally dynamic approach, iterating on this Plan and overall AI efforts to stay ahead of developments and address emerging challenges. This proactive stance will involve continuous benefit and risk assessment, stakeholder engagement, and the implementation of robust safeguards to ensure ethical and equitable AI use. HHS will also continue to identify bold opportunities and collaborations within and across domains that have potential to improve people's lives. HHS divisions will continue to play crucial roles by issuing guidelines and policies, allocating resources, conducting outreach and education programs, and cultivating workforces.

HHS encourages community partners, STLT governments, and other public and private sector partners to responsibly pioneer development and use of AI that improves health and human services for Americans. HHS is committed to collaborating with stakeholders to build on the actions detailed throughout this Strategic Plan and address problems faced in health, human services, and public health, all while ensuring safe and responsible use through the guardrails discussed. HHS will continue to support engagement and transparency with partners to foster creating human-centered solutions with meaningful impact.

As HHS aims to continue its leadership at the forefront of health, human services, and public health innovation to meet the dynamic needs of the American people, this Plan is just one foundational step supporting the Department's ability to address the challenges of tomorrow. HHS is committed to supporting AI that enhances the health and well-being of all Americans.

[817] https://www.whitehouse.gov/briefing-room/blog/2023/12/14/delivering-on-the-promise-of-ai-to-improve-health-outcomes/

[818] https://www.forbes.com/councils/forbesbusinesscouncil/2023/06/16/empowering-individuals-with-disabilities-through-ai-technology/

Appendix A: Glossary of Terms

Table 1: Glossary of Key Terms[819]

Term	Definition
Accountability in AI	The principle that AI systems' creators should be responsible for the outcomes of AI systems, including making amends for any harm caused.
AI ethics	The branch of ethics that examines the moral implications and societal impacts of artificial intelligence.
AI-enabled medical device, AI-enabled device, and/or AI device	In this Plan, the terms "AI/ML-enabled medical device," "AI-enabled device" and "AI device" may be used interchangeably to refer to one or both of (1) AI software that can perform a medical device purpose (e.g., diagnose, cure, mitigate, treat, prevent) without being a part of a traditional hardware medical device; and (2) AI software that is part of or integral to a medical device.
Artificial intelligence (AI)	Per Executive Order 14110, section 3(b), and 15 U.S.C. 9401(3), AI is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action.
Artificial intelligence performance monitoring (AI performance monitoring)	Refers to the process of regularly collecting and analyzing data on the use of a deployed AI system to evaluate its performance in accomplishing its intended tasks in real-world settings. The assessment of an AI model's performance involves various performance metrics and criteria depending on the specific application. This monitoring typically aims to assess the performance of these AI systems in practice, detect performance degradation or changes (e.g., due to data drift), identify instances of misuse, and address any safety or usability concerns.
Artificial intelligence system (AI system)	Any data system, software, hardware, application, tool, or utility that operates in whole or in part using AI.
Assistive artificial intelligence (assistive AI)	AI-enabled products designed to assist human decision-making. The AI only provides suggestions, information, or data that helps users make more informed decisions. Assistive AI and Autonomous AI exist on a spectrum. Examples of Assistive AI might include a wearable device that monitors a patient's vital signs and alerts the user or a healthcare provider when certain metrics are out of the normal range or a product that assists radiologists by showing the location of a potential abnormality.
Autonomous artificial intelligence (autonomous AI)	AI-enabled products that can perform tasks, operate independently, and make decisions without human intervention, such as AI agents. The level of autonomy can vary based on the product. Assistive AI and Autonomous AI exist on a spectrum. An example of Autonomous AI could be a product that autonomously identifies normal X-rays and creates reports without the need for radiologist intervention.
Bias in AI	The introduction of prejudiced assumptions and preferences into AI algorithms and datasets, which can lead to unfair outcomes or decisions.

[819] Definitions sourced from FDA Digital Health and AI Glossary, CMS AI Playbook, and other resources (e.g., government publications or articles).

Term	Definition
Biological product	Per the Public Health Service Act,[820] the term "biological product" means a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, protein, or analogous product, or arsphenamine or derivative of arsphenamine (or any other trivalent organic arsenic compound), applicable to the prevention, treatment, or cure of a disease or condition of human beings.
Chatbot	A program that enables communication between the LLM and the human through text or voice commands in a way that mimics human-to-human conversation.
Clinical decision support (CDS) software	Software that is intended to provide decision support for the diagnosis, treatment, prevention, cure, or mitigation of diseases or other conditions
Cloud computing	A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction
Confabulation in AI	A phenomenon where AI models generate false or misleading information despite being presented with accurate data.
Continual machine learning	The ability of a model to adapt its performance by incorporating new data or experiences over time while retaining prior knowledge/information. The model changes are implemented such that for a given set of inputs, the output may be different before and after the changes are implemented. These changes are typically implemented and validated through a well-defined process that aims at improving performance based on analysis of new data. In contrast to a locked model, a continual machine learning model has a defined learning process to change its behavior.
Convolutional neural network (CNN)	A specialized deep neural network architecture that consists of one or more convolution layers that is suited for processing grid-like data, such as images. In a convolution layer, a “filter” (window or template) slides over regions of the input image to identify low-level patterns (e.g., edges) by applying convolution (a mathematical dot operation applied to the input data). Different filters can be applied to extract different features, such as edges, textures, or curves in images. Additionally, CNNs can include pooling layers, whose function is to reduce the feature dimensionality while retaining relevant features. These convolution and pooling layers get stacked on top of each other to enable this network to build up a hierarchical understanding of patterns and makes CNNs effective at tasks like image recognition and computer vision. An important aspect of this network is its ability to conserve spatial information of the original input while still performing the feature extraction.
Data card	A structured report of relevant characteristics of datasets needed by stakeholders for AI development and evaluation. It contains a descriptive section including descriptive information such as number of samples, collection protocols and associated metadata, and a scorecard section, a quantitative analysis reporting dataset characteristics using relevant criteria and metrics.
Data drift	Refers to the change in the input data distribution a deployed model receives over time, which can cause the model's performance to degrade. This occurs when the properties of the underlying data change. Data drift can affect the accuracy and reliability of predictive models. For example, medical AI-enabled products can experience data drift due to, statistical differences between the data used for model development and data used in clinical operation due to variations between medical practices or context of use between training and clinical use, and changes in patient demographics, disease trends, and data collection methods over time.

[820] https://uscode.house.gov/view.xhtml?req=(title:42%20section:262%20edition:prelim)

Term	Definition
Data governance	The process of managing the availability, usability, integrity, and security of the data in enterprise systems, based on internal data standards and policies that also control data usage.
Data privacy	The aspect of information technology that deals with an organizations or individual's ability to determine what data in a computer system can be shared with third parties.
Data standard	A type of standard, which is an agreed upon approach to allow for consistent measurement, qualification or exchange of an object, process, or unit of information. Data standards refer to methods of organizing, documenting, and formatting data to aid in data aggregation, sharing and reuse.
Data use agreement (DUA)	A legal contract between the entity that owns access to a data source, typically a dataset or database, and a secondary entity that will receive the data, or a subset of it, for reuse. A DUA outlines terms and limitations on how the shared data can be used, and the secondary entity may need to meet certain criteria, such as their affiliated institution, their faculty status, and IRB approval for their research study. Examples of limitations include restricting access to the shared data, requiring that any research dissemination include citation of the data and its originating entity, requiring that data files are destroyed at the completion of research period, and restrictions on data use for commercial purposes. DUAs are frequently required for access to data that contain protected health information (PHI).
Data-driven AI	AI that emphasizes the importance of data in enhancing technology's ability to learn from and augment human intelligence. It involves effective data understanding, governance, and a mindset that extends the value of data toward augmenting business processes through AI.
Deep learning	A specialized branch of ML that involves training neural networks with multiple intermediary (hidden) layers that operate between an input layer that receives data and an output layer that presents the final network output. Each layer learns to transform its input data into a slightly more abstract and composite representation and produces an output that serves as an input for the next layer. As data propagates through successive layers, these models can learn hierarchical feature representations from the input data. For example, in healthcare, deep learning models can be used to identify tumors or suspicious lesions in medical images to support physicians and radiologists in the evaluation of disease.
Deepfake	A video, photo, or audio recording that seems real but has been manipulated with artificial intelligence technologies. The underlying technology can replace faces, manipulate facial expressions, synthesize faces, and synthesize speech. Deepfakes can depict someone appearing to say or do something that they in fact never said or did.
Digital health technology (DHT)	A system that uses computing platforms, connectivity, software, and/or sensors for healthcare and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device. They include technologies intended for use as a medical product, in a medical product, or as an adjunct to other medical products (devices, drugs, and biologics). They may also be used to develop or study medical products.
Digital twin	A set of information constructs that mimics the structure, context, and behavior of a physical asset, is dynamically updated with data from its physical twin throughout its life cycle and informs decisions. The bidirectional interaction between the virtual and the physical is central to the digital twin. Digital twins can enable personalized medicine applications. For example, the digital twin of a patient could inform clinical decisions, such as treatment options and clinical assessments. In addition, digital twins can play a role in assembling large, diverse virtual population cohorts for in silico clinical trials, and in quality assessment and process optimization of drug manufacturing processes.

Term	Definition
Drug	Per the FD&C Act,[821] the term "drug" means (A) articles recognized in the official United States Pharmacopoeia, official Homoeopathic Pharmacopoeia of the United States, or official National Formulary, or any supplement to any of them; and (B) articles intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals; and (C) articles (other than food) intended to affect the structure or any function of the body of man or other animals; and (D) articles intended for use as a component of any article specified in clause (A), (B), or (C).
Ensemble methods	ML techniques that combine multiple models to improve the overall predictive performance compared to using a single model. This involves training a set of base models, such as neural networks, and then aggregating their predictions to make the final prediction. Some common ensemble methods include bagging (i.e., training multiple models on different subsets of the training data and averaging their predictions), boosting (i.e., training models sequentially where each new model focuses on correcting the errors of the previous model), and stacking (i.e., using the predictions of multiple base models as input features for a higher-level “meta-model” that learns how to best combine them).
Explainability	"Refers to a representation of the mechanisms underlying AI systems’ operation." (Source: NIST). Explainability may help overcome the opaqueness of black-box systems (i.e., systems where the internal workings and decision-making processes are not transparent or readily understandable). These explanations can take various forms, including free-text explanations, saliency maps, Shapley Additive Explanations (SHAP), or relevant input examples from data. The primary intent is to answer the question "Why" an AI system made a particular decision. Appropriate Explainable AI (XAI) methods may enable the development of more accurate, fair, interpretable, and transparent AI systems to safely augment human decision-making in healthcare.
Exploratory data analysis (EDA)	An approach to analyzing datasets to summarize their main characteristics, often with visual methods, before making further assumptions or testing hypotheses.
Feature engineering	A ML process where attributes from raw data that best represent the underlying patterns are identified for use in training a specific ML model. It involves selecting, transforming, or creating relevant input variables (known as features) to enhance the performance of ML models. Domain knowledge and data analysis techniques can be used to craft features that capture the inherent relationships in the data. For example, for a model that can predict heart failure, feature engineering on patient data may involve creating a “risk score” by combining relevant features such as age, blood pressure, cholesterol levels, and a history of cardiovascular disease.
Federated learning	A decentralized approach to training ML models. Models are trained by each site on data that are kept locally, and model updates are sent to a central server, whereby the central server aggregates these updates to improve a global model. This method is designed to preserve data privacy, as raw data remain at the local sites and are not centralized. For example, federated learning can allow hospitals to collaborate on a heart disease prediction model without sharing patient data. The model is sent to be trained locally at each hospital, and only the model updates from each hospital, not raw data, are sent back and aggregated. This way, individual patient information remains localized, addressing privacy concerns while still benefiting from a collectively improved model.

[821] https://uscode.house.gov/view.xhtml?req=(title:21%20section:321%20edition:prelim)

Term	Definition
Foundation models	AI models trained using large, typically unlabeled datasets and significant computational resources, that are applicable across a wide range of contexts, including some that the models were not specifically developed and trained for (i.e., emergent capabilities). These models can serve as a foundation upon which further models can be built and adapted for specific uses through further training (i.e., fine-tuning). These models can perform a range of general tasks, such as text synthesis, image manipulation, and audio generation. These models are based on deep learning architectures like transformers and can use either unimodal or multimodal input data.
Generative Adversarial Network (GAN)	A deep learning-based model architecture that normally consists of two competing neural networks, a generator, and a discriminator. The goal of the "generator" is to synthesize fake data to fool the "discriminator", while the "discriminator" tries to discriminate between the synthesized examples (generator's output) and the original training data distribution. The goal of the training is to find a point of equilibrium between the two competing networks, and after the training process, the generator learns to generate new data with the same distribution as the training set. This approach can be used to generate synthetic images.
Generative artificial intelligence (GenAI)	"The class of AI models that emulate the structure and characteristics of input data to generate derived synthetic content. This can include images, videos, audio, text, and other digital content (Source: E.O. 14110). This is usually done by approximating the statistical distribution of the input data. For example, in healthcare, GenAI can be used to generate annotations on synthetic medical data (e.g., image features, text labels) to help expand datasets for training algorithms.
Health information exchange (HIE)	Health Information Exchange allows healthcare professionals and patients to appropriately access and securely share a patient's medical information electronically. There are many healthcare delivery scenarios driving the technology behind the different forms of health information exchange available today.
Human-in-the-loop machine learning	An approach where humans interact with ML models to enhance accuracy and end-user trust in the machine. In human in the loop ML, human interaction is iterative and can lead to continuous performance improvement over time. This interaction is especially relevant in scenarios where the model might be uncertain about its predictions and needs human guidance for verification. Unlike human in the loop ML, supervised machine learning primarily involves human input during the data labeling phase, after which the algorithm trains independently. Labeling or annotation is the process of attaching descriptive information to data. Data itself are unchanged in the annotation process.
Human-centric AI (HCAI)	AI that emphasizes the impact of AI technologies on individuals and society, prioritizing human well-being, needs, and goals.
Interoperability	The ability to communicate and exchange data accurately, effectively, securely, and consistently with different information technology systems, software applications, and networks in various settings, and exchange data such that clinical or operational purpose and meaning of the data are preserved and unaltered.
Key performance indicator (KPI)	A measurable value that demonstrates how effectively an organization is achieving key business objectives.

Term	Definition
Large language model (LLM)	A type of AI model trained on large text datasets to learn the relationships between words in natural language. These models can apply these learned patterns to predict and generate natural language responses to a wide range of inputs or prompts they receive, to conduct tasks like translation, summarization, and question answering. These models are characterized by a vast number of model parameters (i.e., internal learned variables within a trained model). LLMs build on foundational AI models by developing more comprehensive language understanding beyond basic linguistic patterns. For example, in the context of LLMs, chatbot is a program that enables communication between the LLM and the human through text or voice commands in a way that mimics human-to-human conversation.
Locked model	A model that provides the same output each time the same input is applied to it and does not change with use, as its parameters or configuration cannot be updated. In case of AI-enabled medical products, locked models can help ensure consistent performance.
Machine learning (ML)	A set of techniques that can be used to train AI algorithms to improve performance at a task based on data.
Machine learning algorithm (ML algorithm)	Step-by-step procedures or set of instructions followed for performing a task or solving a problem. For example, in ML, algorithms are used to train models using data to solve a specific problem.
Machine learning algorithmic bias (ML algorithmic bias)	The term “bias” is used in various contexts in different fields and industries. In the context of AI, bias refers to the systematic deviation in model predictions or outcomes for certain data points or groups compared to others. Here we are focusing on, algorithmic bias, where such deviations can stem from various sources, such as the characteristics of the training dataset, choices made during model development, data processing irregularities, or biases introduced during data collection or from human decisions. Algorithmic bias can lead to a systematic difference or error in treatment of certain objects, people, or groups in comparison to others, or prediction failures that can result in other risks, where treatment is any kind of action, including perception, observation, representation, prediction, or decision.
Machine learning model (ML model)	A mathematical construct that generates an inference or prediction for input data. This model is the result of an ML algorithm learning from data. Models are trained by algorithms, which are step-by-step procedures used to process data and derive results. AI systems (e.g., AI-enabled medical devices) employ one or more models to achieve their intended purpose.

Term	Definition
Medical device	Per the FD&C Act,[822] "device" means an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is (A) recognized in the official National Formulary, or the United States Pharmacopeia, or any supplement to them, (B) intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or (C) intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes. The term “device” does not include software functions pursuant to section 520(o). Note that some software-based behavioral interventions are medical devices under FDA’s statute, whereas others, such as those software functions that are “intended for maintaining or encouraging a healthy lifestyle” and are “unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition,” are not. See sections 201(h) and 520(o)(1)(B) of the FD&C Act.[823]
Medical products	In this Plan, the term “medical products” refers collectively to drugs, biological products, and medical devices (including some software-based behavioral interventions) as defined in this glossary.
Metrics	Quantitative measures used to track and assess the status of specific processes, projects, or activities.
Model calibration	The process of adjusting predicted probabilities generated by an ML model to ensure that they accurately reflect the observed frequencies of events or outcomes in the real world. For example, if a model is well calibrated and predicts 20% probability of breast cancer for a patient, then the observed frequency of breast cancer should be approximately 20 out of 100 patients that were given such a prediction by the model.
Model card	A structured report of relevant technical characteristics of an AI model and benchmark evaluation results in a variety of conditions, such as across different cultural, demographic, or phenotypic groups and intersectional groups that are relevant to the intended application domains. Model cards also provide information about the context in which models are intended to be used and details of how their performance was assessed.
Model deployment	The process of integrating a machine learning model into an existing production environment to make practical and actionable predictions.
Model fitting	The process of training an ML model to capture underlying patterns in the data by adjusting the training parameters to make the model’s predictions as close as possible to the target values in the training data. This adjustment of the parameters enables the model to generalize its understanding of the data, making it useful for making predictions on new, unseen data. A well-fit model does not overfit or underfit but performs well both on the training data and on new, unseen data, due to correctly capturing the relationships between the input and target variables.

[822] https://uscode.house.gov/view.xhtml?req=(title:21%20section:321%20edition:prelim)

[823] https://uscode.house.gov/view.xhtml?req=(title:21%20section:321%20edition:prelim)

Term	Definition
Model robustness	The ability of an ML model to maintain its target or specified level of performance under different circumstances. These circumstances can include noisy data (e.g., data containing errors, inconsistencies, and missing values), unseen data or data drift, or adversarial attacks that manipulate the data to deceive the model. For example, in healthcare, challenges in model robustness can arise in medical image classification, where variations in imaging conditions like lighting or resolution, can affect the performance of a tumor classification model trained on standardized images.
Model weight	A numerical parameter within an AI model that helps determine the model's outputs in response to inputs.
Multimodal	An approach for processing and integrating multiple different data types, aiming to capture and leverage the relationships between them for a better understanding of the input information or improved prediction performance. These data types may include text, images, audio, video, genomics, sensor data, etc. These different data types may be processed using a single multimodal network (e.g., based on neural network, or other architectures) or through separate unimodal networks (e.g., LLMs for text and CNNs for images) where the unimodal outputs are combined. For example, in healthcare, data from electronic health records and wearable biosensors can be combined to enable remote monitoring of patients.
National Vital Statistics System (NVSS)	The National Vital Statistics System is the oldest and most successful example of inter-governmental data sharing in Public Health and the shared relationships, standards, and procedures form the mechanism by which NCHS collects and disseminates the Nation's official vital statistics. These data are provided through contracts between NCHS and vital registration systems operated in the various jurisdictions legally responsible for the registration of vital events—births, deaths, marriages, divorces, and fetal deaths.
Natural language processing (NLP)	A subfield of AI and linguistics that enables computers to understand, process, interpret, and generate human language. NLP systems can perform tasks such as text classification, sentiment analysis, and translation, using techniques from computational linguistics and ML to process and analyze natural language data. Natural Language Generation is one application of NLP, which involves using AI systems to produce human-readable text outputs like summaries, reports, stories, or responses.
Neural network	A computational model inspired by the structure of the human brain. It is composed of interconnected nodes, or "neurons" organized into layers: an input layer that receives data, one or more hidden layers that process and identify patterns in the data, and an output layer that presents the final network output.
Overfitting	In ML, overfitting occurs when a model learns the training data too thoroughly, capturing not just the fundamental patterns, but also noise or random fluctuations. Such a model might excel on the training data, but struggles to generalize to new, unseen data.
Performance metrics	In the context of AI quantitative or qualitative measures that can be used to assess the ability of a model to produce the desired output for a given task. The choice of the metrics depends on the specific task and the model objectives. Examples of quantitative metrics include accuracy, precision, sensitivity (recall), specificity, F1-score, and Area under the Receiver Operating Characteristic curve (AUC-ROC). Qualitative measures may involve heatmap evaluations or visual interpretations. These metrics enable systematic evaluation, comparison, and refinement of models, and aid in the assessment of whether the model meets its intended objectives.

Term	Definition
Personally identifiable information (PII)	Any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information.
Pharmacovigilance	Per FDA's Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment,[824] which applies to activities with respect to drugs and biological products (excluding blood and blood components), the term "pharmacovigilance" refers to "all scientific and data gathering activities relating to the detection, assessment, and understanding of adverse events. This includes the use of pharmacoepidemiologic studies. These activities are undertaken with the goal of identifying adverse events and understanding, to the extent possible, their nature, frequency, and potential risk factors."
Predictive analytics	The use of data, statistical algorithms, and ML techniques to identify the likelihood of future outcomes based on historical data.
Privacy in AI	The protection of personal data and information in the development and application of AI systems, ensuring data is used ethically and with consent.
Privacy-enhancing technology	Any software or hardware solution, technical process, technique, or other technological means of mitigating privacy risks arising from data processing, including by enhancing predictability, manageability, disassociability, storage, security, and confidentiality. These technological means may include secure multiparty computation, homomorphic encryption, zero-knowledge proofs, federated learning, secure enclaves, differential privacy, and synthetic-data-generation tools.
Proof of concept (PoC)	An early stage of project development that demonstrates the feasibility of an idea or technology to prove its potential application in solving a particular problem.
Protected health information (PHI)	Individually identifiable health information transmitted or maintained by a covered entity or its business associates in any form or medium (45 CFR 160.103). The definition exempts a small number of categories of individually identifiable health information, such as individually identifiable health information found in employment records held by a covered entity in its role as an employer.
Reading comprehension and generation (RAG)	An AI technique used to enhance the understanding and generation of text by providing a data pool for reference, aiming to avoid issues like hallucination in language models.
Reference standard (in artificial intelligence)	The best available method for establishing or measuring the true state or property of the phenomenon being examined, often represented in the form of labeled data in AI. It serves as a benchmark against which the outputs of a model are evaluated. In clinical settings and medical research, a reference standard is a diagnostic measure or method that is the gold standard clinically and is used to validate the results. For instance, a reference standard can indicate the presence, extent, and location of diseases or abnormalities. Labeling or annotation is the process of attaching descriptive information to data. Data itself are unchanged in the annotation process.

[824] https://www.fda.gov/media/71546/download

Term	Definition
Reinforcement learning	A ML approach where a model (or agent) learns by taking actions and getting rewards or penalties through its interactions with an environment. The model learns from the consequences of its actions, rather than from being explicitly taught, and selects its actions based on its past experiences (exploitation) and by making new choices (exploration), which is essentially trial and error learning. For example, in healthcare, reinforcement learning can be used for recommending personalized treatment plans for patients with chronic diseases. The model is given patient data, including their medical history, current health status, and treatment responses, and then suggests a treatment plan. The key is the feedback loop: as patient data is continually updated with information on how well they are responding to the treatment, the model adjusts its recommendations accordingly. This process involves a lot of trial and error, as the model learns from each patient interaction. Over time, through many such interactions, the model becomes more adept at predicting and recommending the most effective treatment plans for individual patients.
Reliability in AI	The ability of AI systems to operate consistently under specific conditions, delivering accurate and dependable outcomes.
Responsible AI (RAI)	AI practices that uphold society's moral values, ensuring AI systems function fairly, as intended, and are accountable for their results. This includes adherence to principles like fairness, transparency, accountability, safety, privacy, and reliability.
Robustness in AI	The strength of an AI system to maintain its performance in the face of changing conditions or when dealing with unexpected or adversarial inputs.
Sandbox	A safe, controlled, restricted environment that allows for testing products, regulatory approaches, and other technologies without being subject to specific regulations that otherwise (i.e., outside the safe, controlled, restricted sandbox environment) wouldn't be allowed by law.
Scalable and interoperable AI	AI that ensures adoption within an organization is efficient, adaptable, and harmonious with existing workstreams, enabling AI-based solutions to grow and operate in sync with the agency's goals.
Self-supervised machine learning	ML algorithms that generate their own labels from the available unlabeled data. Unlike supervised learning, where labeled data are provided, and unsupervised learning, which uncovers hidden patterns without labels, self-supervised learning leverages the inherent structure within the data to create its own labels. This approach is useful when labeled data are limited or unavailable.
Semi-supervised machine learning	ML algorithms that leverage both unsupervised and supervised techniques. Supervised learning techniques are trained using labeled data, while unsupervised learning techniques are trained using unlabeled data. Labeling or annotation is the process of attaching descriptive information to data. Data itself are unchanged in the annotation process. For example, consider the task of diagnosing lung diseases from chest X-rays. A semi-supervised learning model would initially be trained on a small set of labeled X-ray images, where each image has been marked by radiologists as showing signs of specific lung conditions or being normal. The model then uses this knowledge to start making predictions on a larger set of unlabeled images.
Supervised machine learning	ML algorithms where labeled data is provided, and algorithms are trained using the labeled data. Labeling or annotation is the process of attaching descriptive information to data. Data itself is unchanged in the annotation process.

Term	Definition
Synthetic data	Data that have been created artificially (e.g., through statistical modeling, computer simulation) so that new values and/or data elements are generated. Generally, synthetic data are intended to represent the structure, properties and relationships seen in actual patient data, except that they do not contain any real or specific information about individuals. For example, in healthcare, synthetic data are artificial data that are intended to mimic the properties and relationships seen in real patient data. Synthetic data are examples that have been partially or fully generated using computational techniques rather than acquired from a human subject by a physical system.
Test data	These data are used to characterize the performance of an AI system. These data are never shown to the algorithm during training and are used to estimate the AI model's performance after training. Testing is conducted to generate evidence to establish the performance of an AI system before the system is deployed or marketed. For AI-enabled medical products, test data should be independent of data used for training and tuning.
Testbed	A facility or mechanism equipped for conducting rigorous, transparent, and replicable testing of tools and technologies, including AI and privacy-enhancing technologies, to help evaluate the functionality, usability, and performance of those tools or technologies.
Training data	These data are used by the manufacturer of an AI system in procedures and training algorithms to build an AI model, including to define model weights, connections, and components.
Transfer learning	A strategic approach within ML wherein a model developed for a particular task is adapted for a second task. This approach leverages the knowledge and patterns acquired from a previously solved problem (source task) to boost the performance and learning efficiency of a model on a subsequent, often similar, problem (target task). For example, in healthcare, a model trained to identify tumors in lung X-ray images might leverage the learned patterns to improve the identification of abnormalities in liver ultrasound images.
Transparency and explainability in AI	The ability of AI systems to be understood and the processes and outcomes explained in human terms.
Tuning data	This data is typically used by the manufacturer of an AI system to evaluate a small number of trained models. This process involves exploring various aspects, including different architectures or hyperparameters (i.e., parameters used to tune the model for the task). The tuning phase happens before the testing phase of the AI system and is part of the training process. While the AI and ML communities sometimes use the term "validation" to refer to the tuning data and phase, the FDA will not typically use the word "validation" in this context due to its specific regulatory definition (see 21 CFR 820.3(z)).
Underfitting	In ML, underfitting happens when a model does not capture the patterns and complexity of the training data, leading to poor performance on both the training and new, unseen data.
Unsupervised machine learning	ML algorithms that only make use of unlabeled data during training. Unsupervised learning seeks to uncover hidden patterns or structures within the data.
User experience (UX) design	The process of designing products, systems, or services with a focus on the quality and efficiency of the user's interaction with and experience of the product.
User research	Research conducted to understand the behaviors, needs, and motivations of users through observation techniques, task analysis, and other feedback methodologies.
Watermarking	The act of embedding information, which is typically difficult to remove, into outputs created by AI—including into outputs such as photos, videos, audio clips, or text—for the purposes of verifying the authenticity of the output or the identity or characteristics of its provenance, modifications, or conveyance.

Table 2: Acronyms

Term	Full Form Text
ACF	Administration for Children and Families
ACL	Administration for Community Living
AHRQ	Agency for Healthcare Research and Quality
AI	Artificial intelligence
AIDR	AI data readiness
ARPA-H	Advanced Research Projects Agency for Health
ASPR	Administration for Strategic Preparedness and Response
ASTP/ONC	Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology
ATSDR	Agency for Toxic Substances and Disease Registry
CAIO	Chief AI Officer
CBER	Center for Biologics Evaluation and Research
CDC	Centers for Disease Control and Prevention
CDER	Center for Drug Evaluation and Research
CDRH	Center for Devices and Radiological Health
CDS	Clinical Decision Support
CGMP	Current Good Manufacturing Practices
CMS	Centers for Medicare & Medicaid Services
CPT®	Current Procedural Terminology
CRDC	Cancer Research Data Commons
DMI	Data Modernization Initiative
DOE	Department of Energy
ECG	Electrocardiogram
EHR	Electronic health record
FDA	Food and Drug Administration
FTC	Federal Trade Commission
GSA	General Services Administration
HHS	Department of Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act
HRSA	Health Resources and Services Administration

Term	Full Form Text
IDE	Investigational Device Exemption
IHS	Indian Health Service
IND	Investigational New Drug
IRB	Institutional Review Board
LEAP	Leading Edge Acceleration Project
LLM	Large language model
ML	Machine learning
MoA	Mechanism of Action
MRI	Magnetic Resonance Imagine
NCHS	National Center for Health Statistics
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NLP	Natural language processing
NOFO	Notice of Funding Opportunity
NPSD	Network of Patient Safety Databases
NTAP	New Technology Add-on Payment
OCAIO	Office of the Chief Artificial Intelligence Officer
OMB	Office of Management and Budget
PDSI	Predictive Decision Support Interventions
PHI	Protected health information
PI	Predictive intelligence
PII	Personally identifiable information
PoC	Proof of concept
PSO	Patient Safety Organizations
RCM	Revenue cycle management
SAMHSA	Substance Abuse and Mental Health Services Administration
SaMD	Software as a medical device
SDOH	Social determinants of health
TA	Therapeutic area
TPLC	Total product life cycle

Term	Full Form Text
TTS	Technology Transformation Services
UDS	Uniform Design System
XAI	Explainable AI

Appendix B: Select Federal Policies and Regulations

Table 3: Non-exhaustive federal policies and regulations that support responsible use of AI

Policy focus and goals	Specific regulation, policy, or guidance	Brief description
Overarching legislative and executive actions on AI: Lay out coordinated federal approaches on AI broadly, including its implications in the federal government itself, that can improve AI in the U.S. and ensure its continued safe and responsible use.	Executive Order 14110[825] (Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence)	Highlights the importance of enabling continued safe adoption of AI and requires several federal agencies, including HHS, to develop AI strategies.
	Blueprint for the AI Bill of Rights[826]	Identifies five principles that should guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence: safe and effective systems; algorithmic discrimination protections; data privacy; notice and explanation; and human alternatives, consideration, and fallback.
	National AI Initiative Act of 2020[827]	Calls for a coordinated program across the entire Federal government to accelerate AI research and application for the Nation's economic prosperity and national security.
	Executive Order 13859[828] (Maintaining American Leadership in AI)	Defines a coordinated Federal Government AI strategy focused on driving technological breakthroughs, developing appropriate AI standards, training current and future workforces, fostering public trust and confidence, and promoting an international environment that supports American AI research.
	Executive Order 13960[829] (Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government)	Establishes principles for trustworthy AI use in and by federal government agencies.

[825] https:/www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

[826] https://www.whitehouse.gov/ostp/ai-bill-of-rights/

[827] https://www.congress.gov/bill/116th-congress/house-bill/6216

[828] https://www.govinfo.gov/content/pkg/FR-2019-02-14/pdf/2019-02544.pdf

[829] https://www.hhs.gov/programs/topic-sites/ai/statutes/index.html

Policy focus and goals	Specific regulation, policy, or guidance	Brief description
	OMB M-21-06[830] (Guidance for Regulation of Artificial Intelligence Applications)	Provides guidance to all Federal agencies to inform the development of regulatory and non-regulatory approaches regarding technologies and industrial sectors that are empowered or enabled by artificial intelligence (AI) and consider ways to reduce barriers to the development and adoption of AI technologies. HHS responded to this OMB[831] with the statutory authorities that authorize HHS to issue regulations on the development and use of AI applications in the private sector, among additional topics.
	Section 1557 92.210 Nondiscrimination in the use of patient care decision support tools[832]	Protects against discrimination based on race, color, national origin, sex, age or disability in health programs or activities through use of patient decision support tools.
Research Participant Protections: Establish expectations and best practices for protecting the welfare, privacy, and autonomy of research participants. The ethical considerations embedded in these policies, regulations, and best practices (e.g., privacy) address key issues relevant to the development and use of AI in research. In adhering to them, investigators can mitigate potential harms and inequities arising from the use and development of AI.	Protection of Human Subjects (45 CFR 46)[833]	Outlines basic provisions for IRBs, informed consent, and assurance of compliance for HHS-supported research involving human participants and their data, including considerations of risks & benefits.
	Protection of Human Subjects (21 CFR 50)[834] and Institutional Review Boards (21 CFR 56)[835]	Provisions for compliance and IRBs for clinical investigations that are also regulated by FDA.
	Certificates of Confidentiality[836]	Prohibits the disclosure of identifiable, sensitive research information to anyone not connected to the research except when the participant consents or in a few other specific situations.
	NIH Informed Consent for Secondary Research with Data and Biospecimens[837]	Provides points to consider, instructions for use, and optional sample language that is designed for informed consent documents for research studies that include plans to store and share collected data and biospecimens for future use.
	Common Rule[838]	Requires obtaining legally effective informed consent before involving a human subject in research.
	Informed Consent Posting Instructions[839]	Provides general instructions on how to comply with the Common Rule's requirement to gain informed consent before involving human subjects in research.

[830] https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf
[831] https://www.hhs.gov/sites/default/files/department-of-health-and-human-services-omb-m-21-06.pdf
[832] https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-92/subpart-C/section-92.210
[833] https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html
[834] https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=50&showFR=1
[835] https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=56&showFR=1
[836] https://grants.nih.gov/policy-and-compliance/policy-topics/human-subjects/coc
[837] https://osp.od.nih.gov/wp-content/uploads/Informed-Consent-Resource-for-Secondary-Research-with-Data-and-Biospecimens.pdf
[838] https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/revised-common-rule-regulatory-text/index.html#46.116
[839] https://www.hhs.gov/ohrp/regulations-and-policy/informed-consent-posting/informed-consent-posting-guidance/index.html

Policy focus and goals	Specific regulation, policy, or guidance	Brief description
	NIH Information about Protecting Privacy when Sharing Human Research Participant Data[840]	Provides a set of principles and best practices for protecting the privacy of human research participants when sharing data in NIH-supported research. (Issued under the NIH Data Management and Sharing policy.)
Patient Protections: Help protect the privacy and security of health data, including in healthcare delivery, research and discovery, and more.	HIPAA Privacy Rule[841], [842]	HIPAA helps protect the privacy and security of health data used in research, including research involving AI, thereby fostering trust in healthcare research activities. The Privacy Rule establishes the conditions under which protected health information may be used or disclosed by covered entities for research purposes.
	Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule[843]	Implements provisions of the 21st Century Cures Act and makes updates to the ONC Health IT Certification Program (Certification Program) with new and updated standards, implementation specifications, and certification criteria. Provisions in the HTI-1 final rule advance interoperability, improve transparency, and support the access, exchange, and use of electronic health information.
	Health Data, Technology, and Interoperability: Patient Engagement, Information Sharing, and Public Health Interoperability (HTI-2) Proposed Rule[844]	Technology and standards updates that build on the HTI-1 final rule, ranging from the capability to exchange clinical images (e.g., X-rays) to the addition of multifactor authentication support.
	21st Century Cures Act[845]	Helps to accelerate medical product development and bring new innovations and advances to patients who need them faster and more efficiently. The law builds on previous work at FDA incorporating the perspective of patients into the development of drugs, biological products, and devices in FDA's decision-making process and has provisions related to privacy protection and ensuring appropriate access to electronic health information.

[840] https://sharing.nih.gov/data-management-and-sharing-policy/protecting-participant-privacy-when-sharing-scientific-data/principles-and-best-practices-for-protecting-participant-privacy

[841] https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html For the HIPAA Privacy Rule Guidance

[842] https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/combined-regulation-text/index.html For links to the full HIPAA Administrative Simplification Regulations including the Privacy Rule.

[843] https://www.healthit.gov/topic/laws-regulation-and-policy/health-data-technology-and-interoperability-certification-program

[844] https://www.healthit.gov/topic/laws-regulation-and-policy/health-data-technology-and-interoperability-patient-engagement

[845] https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act

Policy focus and goals	Specific regulation, policy, or guidance	Brief description
	Health Information Technology for Economic and Clinical Health (HITECH) Act[846]	Provides HHS with the authority to establish programs to improve healthcare quality, safety, and efficiency through the promotion of health IT, including electronic health records and private and secure electronic health information exchange. The Act addresses privacy and safety concerns related to electronic health information exchange, including with stricter breach notification requirements.
Biosecurity and Biosafety: Establish and are part of a comprehensive biosecurity and biosafety oversight system. Research funded by HHS, including research using the tools and technologies enabled or informed by AI, fall under this oversight framework. While some of these policies do not explicitly address AI, they are still applicable to development and use of AI in research involving biological agents, toxins, or nucleic acid molecules if such research involves physical experiments that are covered under these policies.	U.S. Government Policy for Oversight of Dual Use Research of Concern and Pathogens with Enhanced Pandemic Potential (in effect May 6, 2025)[847]	Provides a unified federal oversight framework for conducting and managing certain types of federally funded life sciences research on biological agents and toxins that have the potential to pose risks to public health, agriculture, food security, economic security, or national security. The policy "encourages institutional oversight of in silico research, regardless of funding source, that could result in the development of potential dual-use computational models directly enabling the design of a [pathogen with enhanced pandemic potential or a novel biological agent or toxin." Once in effect (May 6, 2025), this unified framework will supersede the current oversight delineated through: • USG Policy for oversight of Life Sciences Dual Use Research of Concern[848] and • HHS Framework for Guiding Funding Decisions about Proposed Research Involving Enhanced Potential Pandemic Pathogens[849]
	U.S. Government Framework for Nucleic Acid Synthesis Screening (in effect October 29, 2024)[850, 851]	Encourages providers of synthetic nucleic acids to implement comprehensive, scalable, and verifiable screening mechanisms to prevent misuse of these nucleotides. Builds on earlier guidance from HHS[852] and requires recipients of federal Research and Discovery funds to procure synthetic nucleic acids only from providers that implement these best practices.
	NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules[853]	Establishes safety practices and containment procedures for institutions that receive NIH funding for "basic and clinical research involving recombinant or synthetic nucleic acid molecules, including the creation and use of organisms and viruses containing recombinant or synthetic nucleic acid molecules."

[846] https://www.healthit.gov/sites/default/files/hitech_act_excerpt_from_arra_with_index.pdf
[847] https://www.whitehouse.gov/wp-content/uploads/2024/05/USG-Policy-for-Oversight-of-DURC-and-PEPP.pdf
[848] https://www.phe.gov/s3/dualuse/Documents/us-policy-durc-032812.pdf
[849] https://www.phe.gov/s3/dualuse/Documents/P3CO.pdf
[850] https://www.whitehouse.gov/wp-content/uploads/2024/04/Nucleic-Acid_Synthesis_Screening_Framework.pdf
[851] https://www.whitehouse.gov/ostp/news-updates/2024/04/29/framework-for-nucleic-acid-synthesis-screening/
[852] https://aspr.hhs.gov/legal/synna/Documents/SynNA-Guidance-2023.pdf
[853] https://osp.od.nih.gov/policies/biosafety-and-biosecurity-policy#tab2/

Policy focus and goals	Specific regulation, policy, or guidance	Brief description
Public Access and Data Management and Sharing: Seek to maximize the responsible management and sharing of research products while ensuring that researchers consider how the privacy, rights, and confidentiality of human research participants will be protected. Increasing the availability of data through data sharing allows for more accurate development and use of AI models. These policies help ensure that investigators remain good stewards of data used in or produced by AI models. HHS operating divisions have a robust set of policies aimed at responsible data sharing, including but not limited to, NIH Genomic Data Sharing Policy, NIH Public Access Policy, and NIH Data Management and Sharing Policy.	Public Access Policies[854]	In August of 2022, the Office of Science and Technology Policy released a Public Access Memo[855] directing Federal Agencies with research and development expenditures to make all peer reviewed scholarly publications publicly accessible by December 31, 2025, without an embargo or cost. Additionally, all scientific data underlying these publications must be made freely available and publicly accessible by default at the time of publication. In response, HHS operating divisions have updated and/or developed Public Access Policies to meet this directive (see NIH Public Access Policy[856]).
	NIH Data Management & Sharing (DMS) Policy[857]	Establishes the requirement to submit a DMS Plan and comply with NIH-approved plans. In addition, NIH Institutes, Centers, and Offices can request additional or specific information be included within the plan to support programmatic priorities or to expand the utility of the scientific data generated from the research.
	NIH Genomic Data Sharing Policy[858]	Promotes and facilitates responsible sharing of large-scale genomic data generated with NIH funds.
Licensing, Intellectual Property, & Technology Transfer	US Patent and Trademark Office information about AI[859]	Provides AI-related patent resources and important information concerning AI IP policy.
	NIH Research Tools Policy[860]	Expects funding recipients to appropriately disseminate propagate and allow open access to research tools developed with NIH funding.

[854] https://www.hhs.gov/open/public-access-guiding-principles/index.html
[855] https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf
[856] https://sharing.nih.gov/public-access-policy/public-access-policy-overview
[857] https://sharing.nih.gov/data-management-and-sharing-policy
[858] https://sharing.nih.gov/genomic-data-sharing-policy
[859] https://www.uspto.gov/initiatives/artificial-intelligence
[860] https://sharing.nih.gov/other-sharing-policies/research-tools-policy