U.S. Department of Health and Human Services
Office of the Chief Data Officer

Data Strategy 2023-2028

Table of Contents

Version dated: December 2023

Note: *This strategy is not intended to be comprehensive of all data-related work being done at HHS. These steps are additive to currently active work and intend to further boost HHS' progress in leveraging data in support of its mission. Given the multi-year time frame of this strategy, priority initiatives are subject to review and possible change to maintain alignment with the HHS Data Vision.*

Message from the Chief Data Officer

I am pleased to share the U.S. Department of Health and Human Services (HHS) Data Strategy for 2023-2028. This strategy introduces a bold vision for how data can be leveraged at the Department to meet its mission of ensuring the well-being of Americans.

The Department's previous data strategy was released in 2018. In the subsequent years, many new operational needs and opportunities have emerged that broaden the use of data and possibilities for the Department. HHS established data leadership bodies such as the Office of the Chief Data Officer, Data Governance Board, Data Council, and agency-level Chief Data Officers who can plan, resource, and implement a more comprehensive data strategy. In addition, the COVID-19 pandemic radically altered the state of data at HHS and across the health and human services ecosystem, including forging innovative data sharing partnerships and introducing new tools and capabilities. HHS leveraged and built upon the successful data initiatives of the past to improve the way HHS is delivering services. Finally, recent major developments in advanced analytics and artificial intelligence are rapidly enabling transformative new data-driven approaches. As a result, this strategy focused around five priority areas, where investments are most critical to advance the safe and effective use of data across HHS. These priority areas are: 1. Cultivate Data Talent; 2. Foster Data Sharing; 3. Integrate Administrative Data into Program Operations; 4. Enable Whole-Person Care Delivery by Connecting Human Services Data; and 5. Responsibly Leverage Artificial Intelligence.

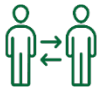
The foundations of this strategy were built upon the input received from HHS mission leadership on how data can be further leveraged to achieve the mission of HHS. To demonstrate the impact of translating data into action this strategy also identified two anchor use cases representative of the highest priority areas where cross-department data action can catalyze delivery on critical mission objectives: (a) the Cancer Moonshot and (b) Preparedness and Incident Response.

Nikolaos Ipiotis
Chief Data Officer (Acting)
Office of the Chief Information Officer
U.S. Department of Health and Human Services

Introduction

The 2023 HHS Data Strategy outlines the Department's priorities and initiatives to safely and effectively harness data to enhance the health and well-being of all Americans. The focus of this strategy is the near-to-medium term, about five years, with an intent to periodically review its content to ensure it is meeting HHS' evolving needs. This strategy is based on the research and recommendations of a cross-HHS task force, convened by the Deputy Secretary in Fall 2022, and composed of data leaders and subject matter experts. This task force assessed the current state of HHS' data use, needs, and capabilities; compiled best practices from within the Department and externally; and developed a list of focus areas where investments are most critical to advance the safe and effective use of data across HHS. Those data priorities are:

Cultivate Data Talent



Foster Data Sharing

Integrate Administrative Data into Program Operations

Enable Whole-Person Care Delivery by Connecting Human Services Data

Responsibly Leverage Artificial Intelligence

In addition, this strategy identifies two anchor use cases representative of the highest priority areas where cross-department data action can catalyze delivery on critical mission objectives: (a) the Cancer Moonshot and (b) Preparedness and Incident Response.

For each priority and each use case, this strategy identifies a 5+ year aspiration, as well as near-term (next 1-2 year) initiatives that can be accomplished within existing resources and authorities. To achieve the full aspiration will require additional time and investment in the out years.

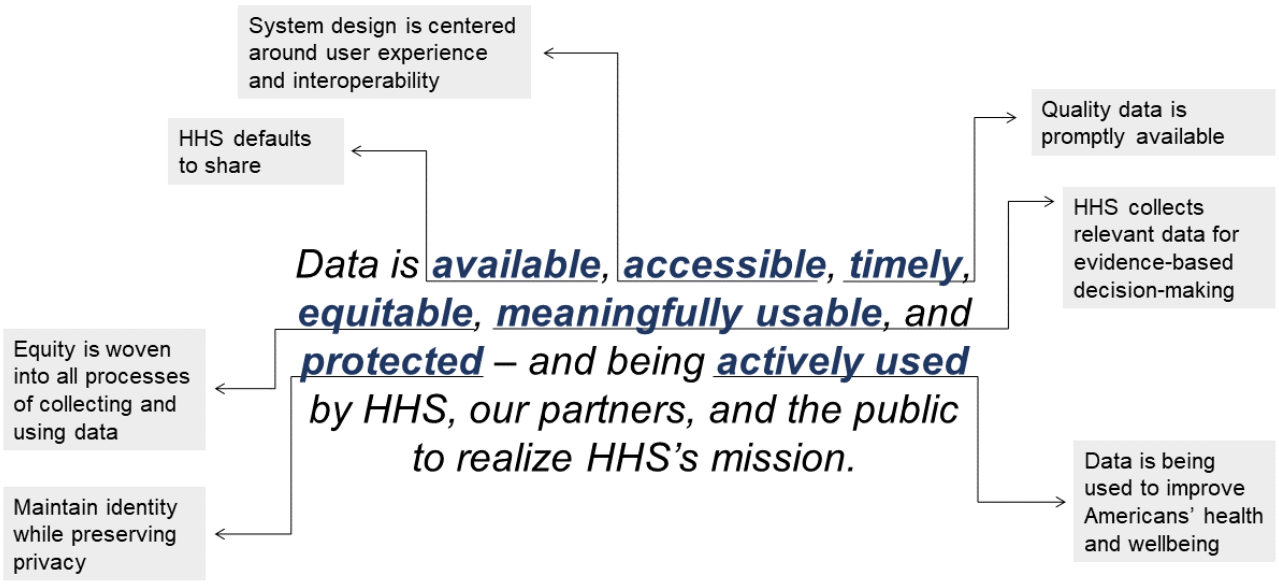
HHS' previous data strategy was drafted in 2018, following the passage of the Evidence Act. That strategy focused heavily on use of statistical evidence to support policymaking and program evaluation. In the subsequent years, many new operational needs and opportunities have emerged that broaden the data mandate and possibility for the Department. HHS has grown its internal data capacity, building on existing cross agency groups, such as the HHS Data Council, and establishing new data leadership bodies such as the HHS Office of the Chief Data Officer, Data Governance Board, and agency-level Chief Data Officers who are able to plan, resource, and implement a more comprehensive data strategy. The COVID-19 pandemic radically altered the state of data at HHS and across the health and human services ecosystem, including forging innovative data sharing partnerships; introducing new tools, models, and personnel; and demonstrating the potential and urgency of data for planning, resource allocation, operations, and coordinated care delivery within the department and in the field. Continued progress in health information exchange, interoperability, and data standards has made new uses of data possible. Recent major developments in advanced analytics and artificial intelligence are rapidly enabling transformative new data-driven approaches. The implementation of these new technologies and innovations may also pose risks to privacy and transparency that must be addressed and mitigated. Considering these new capabilities and needs, the 2023 HHS Data Strategy takes on a broader focus and bolder vision, building on the foundation laid by previous strategies.

Vision

A multi-year strategy cannot anticipate all contingencies and exigencies. When addressing future unknowns, the Department will endeavor to make decisions based on a set of guiding principles, encapsulated in the **HHS Data Vision:**

> *Data is available, accessible, timely, equitable, meaningfully usable, and protected – and being actively used by HHS, our partners, and the public to realize HHS' mission.*

The goal of this strategy is to increase capacity, improve processes and governance, focus efforts and resources, and grow shared infrastructure, while maintaining privacy, to deliver on this vision.



Available

HHS "defaults to share" data that can be disseminated openly or shared within the Department without being asked, as required by the Evidence Act, unless there is a specific rationale why the data cannot be shared. HHS staff are encouraged to collaborate on data projects and make data available for use outside of specific HHS components, while ensuring appropriate privacy and consent protections.

Accessible

HHS designs its systems and interactions with user experience at the center to minimize administrative burden, including for the general public; State, Tribal, Local, Territorial (STLT) governments; delivery partners; researchers; and employees. Data is defined, standardized, catalogued, and interoperable to the extent possible. Processes to access and use data are navigable without significant burden or delay

Timely

Data is promptly available to meet the rapid response needed for care delivery, planning, operations, scientific discoveries, and evaluation. HHS appropriately balances processing time to ensure consistency and quality, with the declining utility of older data.

Equitable

Equity is woven into the processes of collecting and using data, ensuring that a range of populations and lived experiences are accurately reflected in the data. Data and other information are available and analyzed to track progress towards equity and to reduce disparities.

Meaningfully usable

HHS collects data that is relevant to deliver on its mission, and matchable or identifiable when necessary for research and delivery purposes, while preserving privacy. HHS has the technology and people with the right skillsets to leverage data for evidence-based decision-making.

Protected

HHS secures its data to preserve privacy, and ensure confidentiality, integrity, and availability to authorized users.

Being actively used by HHS, its partners, and the public

Data is, in practice, being used by HHS staff; state, local, Tribal, and territorial governments; service providers; researchers; other ecosystem partners; and the general public to improve Americans' health and wellbeing.

Priorities

A comprehensive landscape analysis of HHS' current state and desired future—including a literature review, stakeholder interviews, a data maturity assessment, a best practice catalog, and information collection from all HHS operating divisions—drove selection of five priority areas for focused investment.

1. **Cultivate Data Talent**
2. **Foster Data Sharing**
3. **Integrate Administrative Data into Program Operations**
4. **Enable Whole-Person Care Delivery by Connecting Human Services Data**
5. **Responsibly Leverage Artificial Intelligence**

Priority 1: Cultivate Data Talent

Opportunity

The people of HHS are essential to its vision of making data meaningfully usable and actively used across the Department. Amidst stiff competition from industry for data talent, HHS must improve recruitment, training, mentoring participants in federal fellowship programs, and retention of data professionals including the effective use of partnerships with academic data science programs to ensure the necessary people are in place as a foundation of all other activities in this strategy.

Aspiration

Meet the data workforce needs of HHS, for today and tomorrow.

Initiatives

1.1 *Improve recruitment for data-related positions to meet workforce needs*

Develop a recruitment guide for HHS hiring staff on data-related positions and series, including leveraging special hiring authorities, incentives, job descriptions, and skills-based competency evaluations, to increase competitiveness, efficiency, and promptness of hiring. Ensure Department-wide awareness of these initiatives and authorize HHS components to use recruitment incentives and retention bonuses for data scientists.

Create and circulate an inventory of internship and fellowship programs to recruit early-career data professionals. Review early-career recruitment practices and hiring data to understand and maximize conversion of interns and fellows to permanent hires.

Review current authorities for short-term appointments, including HHS-specific programs previously managed by the Office of the Chief Technology Officer, and all-of-government programs such as Presidential Management Fellows, Presidential Innovation Fellows, CDC Public Health Informatics Fellowship Program, and the U.S. Digital Corps Fellowship Program. Create or expand compelling short-term opportunities for

private-sector data professionals to work on HHS programs and initiatives.

Build partnerships with academic data science programs to strengthen the data science pipeline for new hires.

Create or expand compelling short-term opportunities for private-sector data professionals to work on HHS programs and initiatives.

Build partnerships with academic data science programs to strengthen the data science pipeline for new hires.

1.2 *Upskill staff and increase data science capacity*

Create curriculum and training opportunities for HHS non-technical management on the effective use of data scientists to increase the use and incorporation of data insights into strategic decision-making. Provide training to non-technical staff on data literacy and program evaluation. Build a data curriculum customized to HHS components and job descriptions, specific use cases, and general data literacy, including training for leadership and data scientists.

Build partnerships with academic data science programs to upskill the HHS workforce in data science.

1.3 *Retain qualified staff*

Create conditions in which employees are encouraged and enabled to concentrate on the mission, with minimal administrative burden and maximal support. This includes providing employees with the tools, training, and data access they need to do their jobs effectively.

Priority 2:
Foster Data Sharing

Opportunity

Data sharing within HHS and with partners improves efficiency, unlocks new research and evaluation capacity, enables whole-person care delivery, facilitates effective decision-making, and provides access to critical information to better serve the public. Controls must be in place to ensure data is shared responsibly and within legal limits; all information, storage, and transmission processes must adhere to applicable consent and privacy requirements. However, more and easier data sharing could be unlocked through streamlined processes and the support to navigate complex requirements. Investment in foundational data sharing infrastructure will enable the Department to fully realize the potential of data sharing to further HHS' mission while observing critical protections.

Aspiration

HHS operating and staff divisions provide trusted, high-quality, easily-usable data and metadata across the Department and to external partners.

Initiatives

2.1 *Improve staff awareness of HHS data resources*

Implement mandatory HHS training of relevant staff on data sharing best practices and resources to equip staff with the context needed to understand how data is managed and captured.

Expand current HHS data inventories and catalogs to provide staff a consolidated list of data assets with consistent definitions and streamlined access.

Launch an internal platform (HHS Connect) to consolidate information about data management resources from across the department in a single location. Develop ongoing communications to improve staff awareness of intra-agency data capabilities.

2.2 *Streamline data sharing agreements*

Develop clear, simple, accessible, and division-specific guidelines and templates for establishing data sharing agreements—such as Data Use Agreements (DUAs) and Computer Matching Agreements (CMAs)—and navigating privacy and security requirements.

Streamline the process for data sharing agreements within HHS to encourage collaboration on data-related projects within the Department.

Create a central team to help staff across HHS develop data sharing agreements.

2.3 *Leverage privacy-preserving technologies*

Assess current privacy protections within HHS and develop training materials, best practices, and guidelines to help HHS staff link records across systems while maintaining privacy and security and reducing risk of re-identification.

Identify and invest in improved privacy technologies, such as:

- HHS-wide privacy enhancing technologies and policies to limit the need to share personally identifiable information while permitting data linkages; and
- Zero Trust technologies to protect data and meet federal cybersecurity requirements.

2.4 *Improve the validity and usability of multiple response data*

Improve techniques for collecting and standardizing, to the extent possible, multiple response demographic data (for example, data on sexual orientation/gender identity, multi-racial and/or multi-ethnic identity, or presence of multiple chronic conditions) to improve the validity and utility of these data for use across HHS, including linkage with other data.

Promote the adoption of common data elements for key demographic variables to augment data sharing.

2.5 *Replicate best practices for data modernization projects*

Develop a roadmap for data systems and process modernization, including recommendations, case studies, sample contract language, best practices, and “quick-start” reference materials on how to scope and navigate the process.

Create catalogue of semantic and metadata standards across HHS and encourage adoption of these standards.

Develop guidance for providing support and technical documentation, such as user guides, methodology statements, codebooks, data dictionaries, and underlying data collection tools.

Priority 3: Integrate Administrative Data into Program Operations

Opportunity

Better integration of near real-time and longitudinal administrative data into HHS program operations will enable more effective performance management and data-driven decision making. Beyond programmatic measures, establishing a culture of data literacy and data-driven management practices will help HHS meet the charge of the Evidence Act and other legislated requirements, fulfill HHS' mission, and serve the public.

Aspiration

Leadership and program staff are aware of the benefits of using data to inform their programmatic decisions and internal operations; the high-quality data and analysis tools needed are available; basic data training is available to all staff; excellent data projects are recognized; and the methods used are made available to the rest of HHS. As a result, data is deeply woven into regular program operations and decisions at all levels.

Initiatives

3.1 *Develop and showcase data-driven program management techniques*

Establish a common platform for all HHS staff, including leadership, to access data-related articles, recognize best practices through an annual awards program, and showcase projects that use data in innovative ways to improve operations. Highlighted projects will include tangible blueprints for others to follow.

Facilitate opportunities for HHS staff to observe and celebrate successful data projects and mature data practices to create a robust network of data-curious staff, intended to incentivize participation in similar collaborations.

3.2 *Create data Communities of Practice*

Facilitate the creation, development, and maintenance of HHS Communities of Practice for data professionals that encourage connection, growth, and mentorship amongst members by sharing resources, successes, challenges, and best practices.

3.3 *Create a central team of data professionals to support data projects*

Develop a service delivery model for a central, HHS-wide team of data professionals to assist HHS divisions and offices based on need. Incorporate insights from technical consulting models in other federal agencies, such as the U.S. Digital Service and the General Service Administration's 18F.

Priority 4: Enable Whole-Person Care Delivery by Connecting Human Services Data

Opportunity

Human services lags health care in the availability, interoperability, and electronic exchange of data, particularly at the point of care. There is a need for more comprehensive data standards; freely available, standard taxonomies for human services provider information; better tools to manage consent, map household relationships, and perform matching and linkage; expanded provider onboarding; and aligned incentives and requirements to stimulate data flows. Interoperability improvements would significantly enhance case management, coordinated care and service delivery, closed loop screening and referral with “warm handoffs” between providers, expedited enrollment into benefits and service programs, and enable more efficient and equitable program design and delivery. Collectively, these advances would improve participants’ experience and access to services. The 2009 HITECH Act and CMS’s Meaningful Use program invested tens of billions of dollars in interoperable health data exchange; in the absence of commensurate investment in human services data exchange, progress can nonetheless be made. In light of the extensive health and public health common data standards development in recent decades and the need for cross-sector data sharing to deliver integrated services, human services data must become interoperable with health data, building from common data standards such as the U.S. Core Data for Interoperability (USCDI).

Aspiration

Real-time, relevant data and data-driven insights are available to human services practitioners, health care providers, facilities/programs, and state, Tribal, local, and territorial governments to establish a whole-person and whole-family view of wellness and needs; seamlessly connect people with needed support services; proactively predict future needs and better address them at an individual and system level; effectively manage programs and facilities; evaluate the effectiveness and value of service provision at a local and macro level; and more equitably and efficiently plan and allocate resources.

Initiatives

4.1 *Align investments, requirements, and policy levers*

Identify opportunities to align federal funding, technical assistance opportunities, technology, and program and reporting requirements to promote and support human services data interoperability and infrastructure buildout. Leverage the HHS Health IT Alignment Policy (where applicable).

Coordinate standards and interoperability requirements to be incorporated into human service program contracts and grants as well as relevant regulations that impact information sharing in human services and between health and human services

4.2 *Develop data standards*

Ensure full human services representation in HHS and external standards-development bodies. Identify and prioritize gaps in standards, use cases, infrastructure (i.e., cloud systems, retention), and policy/governance (i.e., consent management, household relationship mapping) based on HHS, ecosystem, and end-user priorities.

Partner with standards setting organizations to strategically develop human services standards, use cases, and tools that address the most critical gaps, including creating USCDI+ data set standards to support human services programs.

Coordinate with federal partners to ensure human services standards and taxonomies are open and accessible to implementers.

4.3 *Launch interoperability pilots*

Develop, implement, and learn from state and community data sharing/data use pilots between health and human services, beginning with scaling established use cases that utilize existing standards, and expanding over time to test new standards and use cases.

Create new data linkages across multiple data sources (i.e., administrative, service records, and electronic health records), including from federally funded institutions serving medically underserved populations, while addressing privacy, consent, and transparency concerns.

4.4 *Designate HHS owner for human services interoperability*

Formally designate the Office of the National Coordinator for Health Information Technology (ONC) with responsibility for leading the development and harmonization of interoperability standards between health and human services.

Priority 5: Responsibly Leverage Artificial Intelligence

Opportunity

HHS has for many years been working in the space of artificial intelligence ("AI") to advance its mission across numerous arenas, including research and discovery, drug and device safety, health care delivery, human services delivery, and public health. HHS engagement with AI to date includes regulatory and educational actions involving the effectiveness and safety of AI-enabled medical devices, transparency and risk management of clinical decision support tools, non-discrimination, and improving research processes.

2023 represents a historically unique moment, calling for leadership in AI. A convergence of factors—such as growth in cloud computing infrastructure improvements, new technologies, and pandemic investments—has accelerated the path to widespread adoption. Available computing power to train and deploy models has grown exponentially over the past three years. America's leadership in global scientific and economic competitiveness hinges on upfront leadership on AI; perfect cannot be the enemy of progress on AI development and deployment.

Further, AI carries potentially disruptive upside potential in health and human services. AI offers the opportunity to rapidly advance scientific discovery, such as the creation of novel vaccines and therapeutics through new chemical formulations or expedited clinical trials by identifying right populations and running synthetic trials on pre-existing compounds. It also can drive increased healthcare system efficiency, particularly reductions in administrative costs, and can improve clinical decision support (rapidly digesting disparate health record and medical journal content to create tailored considerations and recommendations for patient care). Finally, AI can help expedite identification of eligible populations and scale access to available resources, improving health and human services benefit design and delivery. However, AI also presents significant risks—including discrimination, bias, fraud, disinformation, cybersecurity, biosecurity, privacy, and inappropriate uses—that must be thoughtfully addressed to ensure that we are able to reap the benefits of these new technologies and not stifle innovation, while also protecting our core values.

Of note, HHS's priorities are in alignment with and build on its commitments laid out in the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, or EO 14410. As a part of EO 14110, HHS will publish its AI Strategy covering all priorities outlined herein and in the EO by April 2025.

Aspiration

HHS is fully mobilized, seamlessly integrated with, and meaningfully contributing to a whole-of-government-and-industry approach to improving quality, efficiency, access, and outcomes in health and human services through the safe, ethical, and responsible use of AI.

Initiatives

5.1 *Establish policies on AI in health and human services*

Establish an AI Task Force responsible for creating policies and frameworks on responsible use of AI in the health sector (including research and discovery, drug and device safety, health care delivery, and public health) and human services sector, HHS programs, and HHS internal operations policies.

Coordinate with other federal agencies involved in health and human services as appropriate.

Convene experts and interested parties, and identify legally appropriate guidance, standards, and resources to promote the responsible deployment and use of ethical AI-enabled. technologies in health and human services, consistent with federal laws.

Promote compliance with non-discrimination and privacy laws when AI is deployed in health and human services.

5.2 *Advance quality and safety of AI in health applications*

Develop a framework and strategy for the quality assurance of AI-enabled technologies used in healthcare, in collaboration with external partners (potentially including but not limited to academia, industry, and tribal organizations) and consistent with applicable federal laws and regulations.

Develop a framework and strategy for regulating the use of AI or AI-enabled tools in drug-development processes.

Establish an AI safety program in partnership with voluntary, Federally-listed Patient Safety Organizations (PSOs), healthcare providers, and other private stakeholders to monitor and mitigate patient safety events related to the use of AI; leverage insights from the safety program, as appropriate, to inform federal rulemaking, guidance, or other regulatory activities related to AI or AI-enabled tools' use in health care.

5.3 *Leverage HHS funding to advance responsible use of AI in health*

Collaborate with appropriate private-sector organizations to improve the Artificial Intelligence / Machine Learning (AI/ML) readiness and identify and redress bias in HHS supported data and tools.

Prioritize the allocation of Leading Edge Acceleration Projects (LEAP) cooperative agreement awards to initiatives that explore ways to support the development of AI tools for clinical care, real-world-evidence programs, population health, public health, and related research.

Identify and leverage other funding streams across HHS that could benefit from or be harmed by AI, and ensure provisions promoting ethical, safe, and effective use while minimizing risks.

5.4 *Deploy a full range of AI capabilities across HHS*

Establish a set of guidelines for responsible AI use that address ethical, legal, and social implications, including privacy, fairness, transparency, and accountability.

Implement mandatory HHS staff training on responsible AI use to ensure that all AI applications respect data privacy and ethical principles.

Build an AI tools sandbox with a Large Language Model, including development of policies for sandbox use that ensure any experimental projects adhere to HHS guidelines for data privacy and responsible AI use.

Ensure continuous learning across the Department on AI, including knowledge sharing and continuous monitoring for opportunities to improve internal capabilities and assets.

Use Cases

In addition to designating the priorities (crosscutting infrastructure and capabilities that support all corners of HHS in furthering its use of data agnostic of topic), HHS leaders selected two priority use cases: areas where coordinated cross-department data projects around a common goal can enable HHS to achieve core mission commitments. Although data supports nearly all parts of HHS' mission, the selected use cases present the greatest areas of opportunity for progress and urgency of action. The priority use cases are the cancer moonshot and preparedness and incident response.

Use Case A: Cancer Moonshot

Aspiration

HHS has maximized the utility of data to advance the fight against cancer towards the Moonshot goal of reducing the cancer death rate by 50% within 25 years. We have established secure sharing of privacy-protected health data as standard practice throughout research. We are facilitating researchers' sharing and use of available data to achieve rapid progress against cancer.

Initiatives

A.1 *Develop a roadmap and prototype for federated linked data in a secure sandbox*

Create the technical and legal environment for analytical cancer research by developing a roadmap and prototype for a federated linked data sandbox to facilitate the linking of administrative and clinical data for cancer patients while managing privacy and consent. Address data access and use governance, technical design, and security and controls issues.

A.2 *Create and implement data standards and interoperability needed to make prospective health data usable for clinical research*

Define standards for a computable patient health record sufficient for clinical research and evidence development and a process to periodically refresh them.

Develop a computable consent and data provenance collection and tracking mechanism.

Pilot the patient health record and consent/provenance tracking mechanism in prospective clinical trials.

A.3 *Drive innovative public/private data partnerships to expand data access*

Launch cross-government flagship demonstration projects for data infrastructure and clinical evidence generation.

Issue a governmental call to action for health IT developers and health systems to adopt standardized health record data, consent, and data submissions, and participate in new public-private data sharing and research partnerships.

A.4 *Streamline access to HHS linked administrative and clinical data for clinical research*

Streamline funding and HHS researchers' data access for HHS held linked administrative and clinical data to facilitate easier and improved clinical research.

Use Case B: Preparedness & Incident Response

Aspiration

HHS has established a near real-time information management ecosystem for all-hazards situational awareness and response management that tracks hardening and resilience, emerging threats, incident management needs, delivery system status, and response activities. The system spans across health and human services, and down to the state and local level for coordinated ecosystem and cross-government response.

Initiatives

B.1 *Develop an all-hazards data systems blueprint*

Develop a clear roadmap across HHS, which also encompasses external stakeholders, for the data and system capabilities needed and the structure to deliver them. The plan should address technology, data management, legal structures, policy, contracts, program requirements, governance, standards, and staff. In addition, the plan should address the cross-HHS coordinating structures needed for monitoring, detection, and managing response to multiple incident types, **focusing first on severe respiratory illness (COVID+) and natural disasters**. The plan should be informed by learning and data challenges from recent incident responses and tabletop exercises focusing on top emerging threats. The roadmap should address the data sources and access required to support the full spectrum of state / local needs, including health, public health, and human services.

B.2 *Create foundational infrastructure*

Establish consistent data and system standards across preparedness categories (i.e., minimal data necessary, minimum system standards for EHRs, USCDI+). Identify mechanisms to formalize and implement the standards.

Inventory cross-office opportunities to align policies and program, contract, and grant requirements towards a readiness data framework.

Streamline and expand use of inter- and intra-governmental Data Use Agreements / Memoranda of Understanding to facilitate multi-lateral data exchange, while maintaining privacy and securing data