

Combined Data Analyses

Goal Leaders: Mona Siddiqui, Office of the CTO, APG Goal Leader
Rashida Dorsey, ASPE, Deputy Goal Leader

Overview

Goal Statement

- Increase combined analysis of disparate datasets in order to achieve better insights. By September 30, 2019 HHS will develop and implement an enterprise wide data governance model. This will enable more efficient and effective processes for sharing inter-agency data beyond a dataset's primary purpose. Two use cases of inter-agency data sharing through this process will be tested.

Challenge

- Siloes - HHS has rich data assets that are collected and maintained by separate agencies within the Department. In many instances, Operating Divisions may have specific authorities or mandates to collect data making joint collection or integration of data collection activities difficult.
- Statistical agencies and units must adhere to legal provisions (e.g. CIPSEA) that can inhibit data sharing. HHS programs and services can also be restricted from sharing data by legal provisions. For example, the National Database of New Hires (NDNH) is a valuable data asset, but it contains the highest degrees of protection.
- CMS, IHS and other Health Insurance Portability and Accountability Act (HIPAA) covered entities must adhere HIPAA regulations for the disclosure and use of covered health information
- Agencies within HHS each serve specific missions and developed systems overtime to address their unique needs. However, reexamining the underlying infrastructure can help identify gaps, areas for improvement, and efficiencies. Specifically, HHS needs to develop and implement an enterprise wide data governance strategy.

Overview

Opportunity

- Multi-sector issues—when an issue touches more than one agency or crosses into multi-disciplines than joint research and analyses occur and benefit everyone involved.
- Technology--advances in computing and IT infrastructure allow for secure sharing of information and enable collaborative work.
- Methodology—advances in statistics, machine learning, data mining, and geographic analysis enable disparate datasets to be combined to reveal new insights and find patterns.

Leadership

Core Team

- Mona Siddiqui, APG Goal Leader
- Rashida Dorsey, Deputy Goal Leader
- Reimagine HHS Leveraging the Power of HHS Data Core Team

Goal Structure & Strategies

- HHS agencies have experience conducting combined analyses and linking datasets together (both internal and external). However, these are often treated as “one-offs”, experimental, or exploratory rather than a tool to support the overall mission. Encouraging agency collaboration to conduct this work will require providing tools, devoting staff time, and establishing a framework that benefits all involved.
- An HHS wide data governance plan will establish ownership and rules for data assets and improve data quality and uniformity.
- A streamlined data use agreement infrastructure will enable more seamless interagency data sharing
- Developing an infrastructure that allows for secure and simple sharing of data assets will enable agencies and staff to collaborate on issues that affect multiple agencies such as the opioid epidemic.
- Comparing commonly used data architectures and data management tools to open source software and file formats could reveal more efficient ways to analyze and manipulate data assets. Adopting and using flexible open source solutions would also increase the Department’s ability to partner with the private sector.

Summary of Progress – FY18 Q2

- HHS has begun the process to develop an enterprise-wide data management strategy, with a focus on data governance and data sharing for the Department.
- The enterprise wide data management strategy is being developed in several Phases.
- Phase I included structured interviews with leadership and staff from HHS Operating Divisions and Staff Divisions, representing 30-40 high value datasets.
- Phase I interviews were designed to gain insights on data sharing and governance from data asset owners. This included assessing the functional and technical capacity to share data.
- Phase I interviews are 100% complete

Key Milestones

Milestone Summary					
Key Milestone	Milestone Due Date	Milestone Status	Change from last quarter	Owner	Comments
Develop framework for Department-wide Data Governance Plan	End of Q3, 2018				
Deploy two pilots to test the process and infrastructure	Q4, 2018				
Implement Enterprise wide data governance model	Begin in Q4, 2018				
Development of Department-wide Data Use Agreements (DUAs)	Q4, 2019				