Combined Data Analyses

Goal Leader: Mona Siddiqui, Office of the CTO

Deputy Goal Leader: Amanda Cash, ASPE

Overview

Goal Statement

- o 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.

Challenges

- o 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.

- o Statistical agencies and units must adhere to legal provisions (e.g., Confidential Information Protection and Statistical Efficiency Act) that can inhibit data sharing. HHS programs and services can also be restricted from sharing data by legal provisions. For example, the National Directory of New Hires (NDNH) is a valuable data asset, but it contains the highest degrees of protection.

- o 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

- o 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

- o 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.

- o Technology--advances in computing and IT infrastructure allow for secure sharing of information and enable collaborative work.

- o 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

- APG Goal Leader: Mona Siddiqui, Chief Data Officer, Office of the CTO
- Deputy Goal Leader: Amanda Cash, Deputy Goal Leader, ASPE
- Reimagine HHS: Leveraging the Power of HHS Data Core Team

Goal Structure & Strategies

o 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.

o An HHS wide data governance plan will establish ownership and rules for data assets and improve data quality and uniformity.

o A streamlined data use agreement infrastructure will enable more seamless interagency data sharing.

o 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.

o 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 Q4

o 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.

o The enterprise wide data management strategy is being developed in several Phases.

o 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 report is nearing completion.

o Phase II was focused on defining the future state of data sharing at HHS. We identified 5 agencies (CMS, CDC, AHRQ, HRSA and SAMHSA) that will partnered with us for this stage.

o Agency interviews have been completed.

o These agency stakeholders have worked with the Reimagine Data Team to identify technical requirements for a data sharing platform.

o The CIO team that will help in developing this platform has been identified and is now working with the Reimagine Data Team.

o The team presented to the Executive Review Board (ERB) for stage gate 1 review the first week of October. The ERB provides approval for proceeding with the proof of concept Enterprise Performance Life Cycle process. The team obtained approval with a condition of providing a timeline of activities. This is the first step in exploring the feasibility of a data sharing platform. ERB is a parallel but distinct process and will be instrumental in deploying Phase III of this project.

o The "Future State" report is completed and under internal Department review.

o The Deputy Secretary has been briefed on the report.

o Phase III will begin the implementation of the governance and technical platform through proof of concepts focused on data sharing for understanding the opioid epidemic. The lessons learned will help delineate the technical requirements for future actions.

o Prior to initiating Phase III, agency stakeholder engagement sessions will be held to delineate engagement strategy moving forward.

Key Milestones

Milestone Summary					
Key Milestone	**Milestone Due Date***	**Milestone Status** *[e.g., Complete, On-Track, Missed]*	**Change from last quarter** *[optional column]*	**Owner** *[optional column]*	**Comments** *[Provide discussion of Progress, changes from last update, Anticipated Barriers or other Issues Related to Milestone Completion]*
Develop framework for Department-wide Data Governance Plan	Q4, FY 2018	Complete			The framework for the Department-wide Data Governance Plan was completed by the end of October.
Deploy two pilots to test the process and infrastructure	Q1, FY 2019	On track			The Goal Team is finalizing plans to launch two separate pilots in Q1 to test the process and infrastructure.
Implement Enterprise wide data governance model	Q1, FY 2019	On track			The Goal Team is developing a plan for applying information gathered during the pilots.
Development of Department-wide Data Use Agreements (DUAs)	Q4, FY 2019	On track			The Goal Team is working with partners across HHS to identify elements to include in DUAs.

*Previous Action Plan documentation described expected APG milestone progress to be completed during quarters of the Calendar Year, rather than the Fiscal Year. For the Q4 update, the Key Milestones Table has been converted to reflect APG progress during quarters of the Fiscal Year.