Health and Human Services’ Quest to Becoming a Data-Driven Organization

A new report shows HHS still faces technical, legal and cultural barriers to seamless, internal data sharing.

Amanda Ziadeh
Fri, 09/21/2018 - 17:47

The Health and Human Services Department plans to implement an enterprise wide data strategy to overcome its ultimate challenge: a lack of consistent and standard protocols for interagency data sharing.
But that’s not the department’s only barrier, according to a recent report on the state of data sharing at HHS. A team led by Mona Siddiqui, its chief data officer, interviewed agency leadership from 11 HHS agencies about the challenges of data sharing, focusing on high value, restricted and nonpublic datasets.

Because, in order for HHS to become a data-driven department, it needs a cohesive data governance structure, a platform open to data sharing and policy designed around evidence. Plus, there’s potential for machine learning and artificial intelligence to solve complex health challenges with that data.

“There’s an incredible amount of lag in the information that we often get, so there could be a lag of a year-and-a-half to two years on the data that we get,” said Siddiqui in an interview with GovernmentCIO Media. “To be able to make policy-relevant decisions in time . . . what we really need is access to as close as real time information as possible.”

HHS has seen small wins on a project-by-project basis, but data is largely kept in silos within its agencies, with little awareness of what data is actually collected and how to access it. The report identified five core data challenges restricting internal data sharing, and opportunities for improvement:

**Process for Data Access**

HHS lacks a consistent, transparent and standardized framework for sharing restricted and nonpublic data from agency-to-agency, and the process for requesting access is lengthy and inefficient. In fact, each agency can interpret the rules for its data sharing process autonomously, and agencies aren’t accountable for how they respond to requests for access — even if it’s inappropriately denied or if access is delayed. And when data is shared, various forms to document the exchange are used.

This was surprising to Siddiqui.
“Requesting data from another agency, that entire workflow of managing the request itself, which sounds incredibly boring, but that’s a real lift for a lot of the agencies,” she said, “especially the smaller ones.” Tracking all the requests and documents should be a seamless process.

**Technology for Data Access and Analysis Varies**

The technical forms used for sharing data across agencies vary, the analytical tools used to interpret that data can be redundant, and while agencies are tracking who has access to restricted data, it’s an inconsistent process and they can be challenged in audits for misuse or misinterpretation.

Machine-readable file formats, for example, are used but aren’t yet the default enterprisewide, posing a labor-intensive step and challenge for data analysts.

There’s also areas where HHS can consolidate redundancies on software and acquisition of data tools.

“The acquisition of technology is often done agency by agency, so it’s hard to have oversight of everything that’s happening,” Siddiqui said, but rather than technology itself, it’s about understanding agency needs, and making sure HHS has the resources designed to address those needs.

**Regulatory Environment Challenges**

Each agency’s data collection effort has its own statutes, regulations and policies that govern the collection of and access to the data, and those statutes vary in limitations.

Some statutes specify who exactly can access the data and for what purposes, and others govern data collection efforts across multiple data collection efforts and agencies. This process also includes state agreements and the variety of timeliness and granularity of how states and federal agencies share data, and the interagency agreement process which requires two different legal agreements when agencies share unpublic data. Agencies describe these processes as time-consuming, confusing and laborious.
Disclosure Risk Management

Agencies are concerned about maintaining confidentiality and protecting personally identifiable information, especially as the risk of violating privacy increases as more variables and granular data are collected and shared. Overall, this is leading to an increase on the limits on microdata access, but these restrictions are creating challenges to sharing this data.

PII is either not collected or scrubbed before datasets are shared, which means agencies on the receiving end aren’t getting any datasets that are more unique than what’s public. These laws, consent forms and governance can limit the amount of microdata shared, and the process to request use of the data for something other than its intent is a rigorous one.

Norms and Resource Constraints

“We talk a lot about data sharing, but we’re not resourced appropriately across our agencies to do that,” Siddiqui said. “Having staff that are dedicated to facilitating data sharing and having staff that can help and support any efforts used to be able to do that, that’s been a real challenge.”

Some data reps don’t see the demand for sharing nonpublic data and view public files as enough for analysis. But HHS is faced with strained resources, a culture of fear of misrepresentation of the data, and a reluctance to critique other agencies for unsatisfactory data sharing practices, according to the report. Many people interviewed didn’t even know what data existed across agencies.

Siddiqui found the cultural barriers to data sharing to be one of the most important challenges to tackle. “I think changing the way people are work everyday, and the way people approach data sharing, that takes some time and that takes demonstrating value from what it is we’re trying to engage them in,” she said.

And a fear of misrepresentation of data is a real concern, as it “can have policy implications,” Siddiqui said, so it’s important to provide support in this area as well.
So, What’s Next?

While HHS is working on developing processes to enable data sharing and addressing technical needs, it also needs to engage in real collaboration and change management with agency partners, and demonstrate value from using data, Siddiqui said.

And, each challenge shouldn’t be tackled in isolation, but rather, each incrementally, to “work at developing and implementing them slowly over time,” said added, with feedback and adjustments as needed.

Plus, many agencies within HHS are developing their own internal data strategies while HHS develops an enterprisewide approach, and Siddiqui said those will all eventually harmonize.

But understanding the HHS data landscape and the opportunities was the first step of creating this enterprisewide data sharing framework. The second phase is defining what the future state for data sharing should look like, and HHS is working with agency partners to define what they would like to see in a solution, according to Siddiqui.

Then, phase three is taking an agile approach to building a governance structure and tool to make data sharing possible. Siddiqui said HHS is currently in planning stages for this, and hopes to deploy a solution by the end of the year.

Outside organizations can help, too. “What would be helpful for us moving forward is to get the same sort of information sharing back from other organizations, so that we don’t reinvent what has already worked in other places,” Siddiqui said. “We would love to learn from large complex organizations that have done similar endeavors.”

In the end, HHS’ ultimate goal is to have a policy strategy informed by data.