

Health Data Initiative Progress

Damon L. Davis

U.S. Department of Health & Human Services,
Office of the Secretary,
CTO's Office-IDEA Lab
Innovation, Design, Entrepreneurship & Action





To cap the work of the past three years, we have re-evaluated what we have done and set just released the **1st open data execution plan**.

The plan has 5 goals of which are:

Advancing the [HealthData.gov](https://www.healthdata.gov) site to a more efficient, user friendly, technically advanced platform for data discovery.

Highlighting departmental assets that support **achieving HHS strategic initiatives** and an increased focus on **strategic data liberation**.

Educating new and existing, internal and external participants in the health and human services ecosystem about **data availability for innovative applications, and disseminating the data for problem solving**.

Enabling and incentivizing the health data ecosystem to utilize all data assets in innovative ways.

Implementing **administration and departmental policies** that foster openness and transparency.

Everything I discuss today in some way **leads back to these 5 essential goals** around open data.

Advancing the HealthData.gov site

- Improving metadata quality for better usability - **In Progress (-)**
- Area of the platform for non-HHS data – **In Progress (+)**
- User feedback loops for better usability – **In Progress**
- Developing and publicizing the ability to store and host HHS data – **In Progress**
- Improve sort and search – **Not Started**

- Improving metadata quality for better usability - **In Progress (-)**
 - Updated the online template for manual catalog entries to support Data.json file requirements
 - This task needs more work
- Area of the platform for non-HHS data – **In Progress**
 - Nine states on HD.G – IL, NY, MO, HI, MD, CO, OK, WA, OR
 - TX, FL, CA in progress
 - Soon UK data on obesity
 - Begun outreach to large cities – Austin, Seattle
 - USDA federating data – working on Veterans Affairs
 - talking to CFPB re cut of Debt collection complaints in the health domain
 - Working to federate VA data
- User feedback loops for better usability – **In Progress**
 - Quarterly HD Leads meetings used to intro functionality and take feedback on platform
 - Individual HD leads meetings are feedback opportunity
 - Pilot functionalities with small user groups to refine new functionality before larger roll out
- Developing and publicizing the ability to store and host HHS data – **In Progress**
 - Constant drumbeat of “we could host the data” – CPI, ACF, CMS and more have all been directly offered – few takers
- Improve sort and search – **Not Started**
 - **Will be folded into a larger HD.g refresh**

Highlighting departmental assets that support achieving HHS strategic goals

- Draft a working definition for “Strategically Relevant Data Assets” (SRDA) and generate a list of those HHS data assets - **In Progress**
- Further drive greater data openness – **In Progress**
- Work with each HHS division on external outreach to broad as well as targeted specific ecosystem participants - **Not Started**
- Increase the traffic to [Healthdata.gov](https://healthdata.gov) as the discovery zone for federal & state, and local health & human services data – **In Progress(+)**

- Draft a working definition for “Strategically Relevant Data Assets” (SRDA) and generate a list of those HHS data assets – **In Progress (-)**
 - Consulted some HD Leads on an approach
 - Next step to put high valued datasets into HHS Strat Plan categories for HD Leads to review and suggest more
 - Suggested we search web logs for recurring keywords to inform strategic relevance
 - Need data of frequency of traffic directed from HD.g to OPDIV sites
 - Further drive greater data openness – **In Progress**
 - Lots of meetings with and outreach to non HD Leads across all divisions to explore and encourage new data postings and communications from CPI, CMMI, ASPR, ACF
 - Posting updates to the HDI Yammer groups and All company about HDI activities, admittedly infrequently
 - Work with each HHS division on external outreach to broad as well as targeted specific ecosystem participants - **Not Started**
 - Increase the traffic to [Healthdata.gov](https://healthdata.gov) as the discovery zone for federal & state, and local health & human services data – **In Progress (+)**
 - Increased blog content has improved readership numbers
 - No evidence, anecdotal or otherwise, about downloads of data – need to work on that with the web teams from various divisions.
- ~~

Help divisions to recognize opportunities in data openness and transparency (e.g., transforming frequent Freedom of Information Act (FOIA) requests into open and transparent datasets that eliminate repetitive FOIA responses).

Educating new and existing, internal and external participants

- Spotighting the value of openly available health data to health care transformations – **In Progress**
- Increasing the percentage of machine readable data assets – **In Progress**
- Continuing and expanding external outreach – **In Progress+**
- Develop use cases and an internal marketing to enhance HHS workforce engagement to continue the culture shift toward an understanding of and appreciation for open data – **Not Started**
- Insert the tenets of the HDI into IDEA Lab programs – **In Progress**

- Spotighting the value of openly available health data to health care transformations (CMMI) -
- Increasing the percentage of machine readable data assets – **In Progress**
 - All HD Leads are repeatedly asked to consider their annual publications as open data and machine readable where applicable
 - HIW is working on APIs for datasets
 - Open FDA
 - BD2K is sure to lead increased machine readable data resources
- Continuing and expanding external outreach – **In Progress**
 - One Codecademy posted by NCHS Jim Craver
 - Exceeding our 1 blog/wk goal
 - Multiple speaking engagements already in 2014
- Develop use cases and an internal marketing to enhance HHS workforce engagement to continue the culture shift toward an understanding of and appreciation for open data – **Not Started**
- Insert the tenets of the HDI into IDEA Lab programs – **In Progress**

Enabling and incentivizing the health data ecosystem

- Continue to publicize the availability of the data and draw attention to uses of the data in the transformation of health, healthcare, and the delivery of human services – **In Progress**
- Seek new ways to engage entrepreneurs who may use the data as fuel for their business – **Not Started**
- Develop relationships and support the needs of federal and non-federal data projects, data enclaves, repositories, and innovation test beds for more powerful analytic capabilities – **In Progress**

- Continue to publicize the availability of the data and draw attention to uses of the data as one of the biggest drivers in the transformation of health, healthcare, and the delivery of human services. This includes highlighting the ample opportunities for innovation in the healthcare ecosystem that can be supported by the vast stores of data. – **In Progress**
 - Accepting several meetings per week from innovators, directing them toward data resources, and connecting them to OPDIV POCs for further information gathering
 - Plans for API class with 18F, project with Microsoft Power BI,
- Seek new ways to engage entrepreneurs who may use the data as fuel for their business. – **Not Started**
 - a. Develop and cultivate private sector and community relationships.
 - b. Provide insights about the data to the business community to highlight the myriad opportunities the data provide.
 - c. Generate market analysis for the department about how the data are being used, what additional data the market desires or requires, and accumulate recommendations that will help prioritize the HDI's work.
 - d. Suggested we cultivate relationships with Health Accelerators to learn entrepreneurial needs & establish relationships
 - e. Need to coordinate with ONC's health innovation agenda
- Develop relationships and support the needs of federal and non-federal data projects, data enclaves, repositories, and innovation test beds to meet the demand for more powerful analytic capabilities. – **In Progress**
 - CMS data sharing progressing

Implementing administration and departmental policies that foster openness

- HDI Charter to institutionalize the Health Data Initiative - **Complete**
- Implement the Open Data Policy M13-13 – Managing Information as an Asset – **Complete (-)**
- Draft plans to address the impacts of the Holdren Memo on Increasing Access to the Results of Federally Funded Scientific Research – **In Progress**
- Clarifying the relationship of the Open Data Policy, Holdren Memo, and Executive Order for the implications of each individually and as contrasted with each other – **In Progress**

•Draft an HDI Charter to institutionalize the Health Data Initiative as a vehicle for transparency, public access to health data, and feedback to the department's senior leadership. The charter will clearly articulate a purpose, scope, and goals for the initiative. – Complete
 •Charter Signed May 2014

•Implement the Open Data Policy–Managing Information as an Asset (May 9, 2013) which calls for the collection or creation of information in a way that supports downstream information processing and dissemination activities and must incorporate the following requirements: build information systems to support interoperability and information accessibility; strengthen data management and release practices; strengthen measures to ensure that privacy and confidentiality are fully protected and that data are properly secured; incorporate new interoperability and openness requirements into core agency processes.

•Since the policy was largely based on activities we've completed this was marked Complete with the understanding that we're in maintenance mode with policy updates as they develop.

•Draft plans to address the impacts of the Holdren Memo on Increasing Access to the Results of Federally Funded Scientific Research (February 22, 2013) which requires Federal agencies with more than \$100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and requiring researchers to better account for and manage the digital data resulting from federally funded scientific research. Elements of the requirement include ensuring the public can read, download and analyze in digital form final peer-reviewed manuscripts or final published documents; facilitating easy public search, analysis and access to peer-reviewed pubs from fed-funded research; and ensuring full access to publications' metadata. The HDI activities will support the strategy development and implementation plans drafted by AHRQ, NIH, CDC, ASPR. – **In Progress**

•Ongoing discussions and coordination efforts are in progress with research agencies to identify gaps, opportunities, lend clarity, and generally coordinate the department's response.

•Continue to be a leading agency in the execution of the Executive Order on Making Open and Machine Readable the New Default for Government Information (May 9, 2013) which states that the default state of new and modernized Government information resources shall be open and machine readable; government information shall be managed as an asset throughout its lifecycle to promote interoperability and openness; and agencies shall ensure that they safeguard the individual privacy, confidentiality and national security – **In Progress**

•Open data activities continue to mature at the department but the message that outputs of recurring reports, current procurements, grants, research and other activities should be open data is penetrating.

•Distribute across the department procurement and acquisition language supporting the integration of open data requirements into Federal acquisition and grant-making processes (For civilian agencies, procedures for utilizing these clauses appear within FAR subpart 27.4 – Rights in Data and Copyrights (48 C.F.R. 27.400 et. seq.,)3) – **In Progress**

•Working with contracting and procurement offices in HHS and OMB on language and procedures that support altering new and existing activities to support open data

•Clarifying relationship of the Open Data Policy, the Holdren Memo, and the Executive Order for the implications of each individually and as contrasted with each other – **In Progress**

•Many policies are still being clarified and new policies such as the Administrative Data policy have been introduced. As the policy landscape evolves, these clarifications will continue.

Other Considerations

- Timeline for the Strat & Ex plan was tight, thus slippage on date driven goals
- However some goals are further along than planned
 - Data federation
- The time and resources of other participants are factors
 - 6 educational modules depends on others – need to locate existing activities or develop new ones
- Lots of shiny objects
 - Data viz with OBMT
 - Dashboard for public reporting
- Need to accumulate methodologies for some objectives
 - Metadata improvement
 - Increasing machine readable formats

IDEALAB

INNOVATION · DESIGN · ENTREPRENEURSHIP · ACTION @ HHS