

HealthData.gov Report: Draft Recommendations

NCVHS Work Group on HHS Data Use and Access

Presented by:

Erika Martin

February 23, 2017

Guiding vision and mission

- **Vision statement:** “driving economic empowerment, innovation, and transformation in health”
- **Mission statement:** “HealthData.gov facilitates the transformation of public data into health innovation by powering visibility, competition, community learning, research, and evidence-based decision making.”

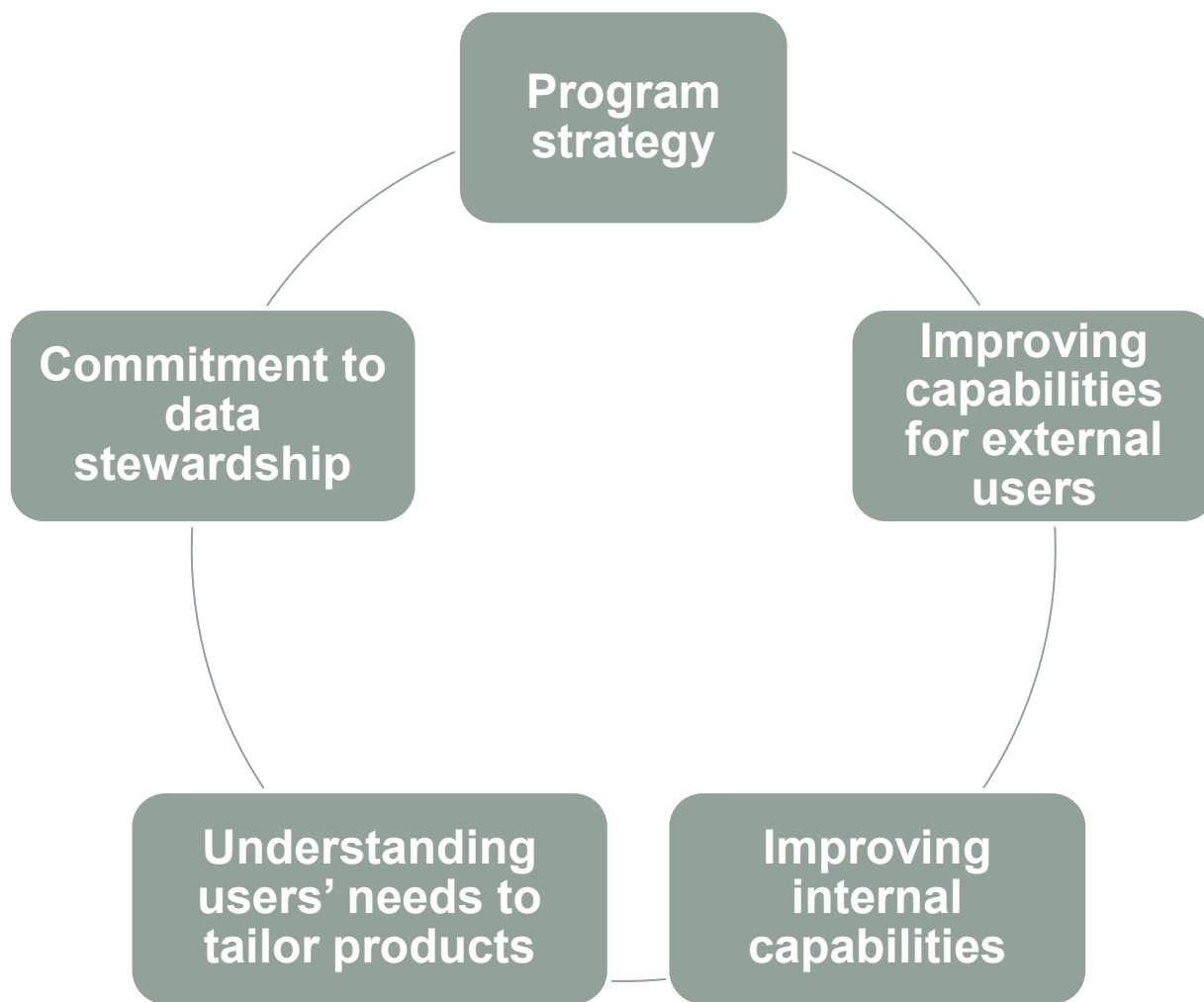
Context and purpose of recommendations

- Context
 - Completed transition from initial developmental phase to covering all HHS operating divisions and multiple state and local agencies
 - Development of a nascent open data ecosystem
 - Organic evolution
 - HealthData.gov has already achieved important outcomes related to the vision and mission
- Purpose of recommendations
 - Enable HealthData.gov to reach its full potential
 - Ensure HealthData.gov is sustainable in the long-term
 - Provide strategies to make future activities more strategic

Support for recommendations

- Discussions among work group, with diverse expertise
 - Best practices in the field (e.g. entrepreneurs, other users)
 - Ideas from other concurrent NCVHS activities
 - “Wish lists” from outgoing and current HealthData.gov leadership
- Peer-reviewed and other literature
 - Mostly focused on New York and other early innovators
 - Gray literature describing use cases (e.g. news articles, white papers)
- Conversations with leadership of the New York and California open health data portals
 - Focused on recommendations

Overview of recommendations



Five recommendations

1. Develop an integrated and coordinated strategy with HHS operating divisions, states, and local municipalities to advance the HealthData.gov mission and vision.
2. Enhance the platform's capabilities to make the data more meaningful to a range of data customers and extend its reach.
3. Develop a long-term HealthData.gov infrastructure that allows it to tap into state-of-the-art knowledge about big data, machine learning, and algorithms for increasing customer use.
4. Implement mechanisms to solicit data customers' input regularly to optimize their use of HealthData.gov and development of data-driven health innovations.
5. Implement data stewardship practices to keep the commitment to the public to protect health data security and confidentiality.

1. Develop an integrated and coordinated strategy with HHS operating divisions, states, and local municipalities to advance the HealthData.gov mission and vision

- Promote HealthData.gov as a brand, reinforcing a consistent message about its vision and mission, and add a tag line such as “Public Data Unleashed” to materials to make the website recognizable and memorable.
- Facilitate a learning collaborative with state and local municipalities to share lessons, such as sustainable business models for public data platforms.
- Inspire data customers to do more with the data by highlighting success stories, encouraging HHS data custodians to post blogs and other reports about their data on HealthData.gov, and facilitating user group meetings.

1. Develop an integrated and coordinated strategy with HHS operating divisions, states, and local municipalities to advance the HealthData.gov mission and vision

- Launch a public data accelerator to increase the number of innovators and data entrepreneurs and successful American businesses using public data.
- Support developer challenge competitions; sponsor events such as the Health Datapalooza; and provide seed funding opportunities to support solutions that make government more efficient and effective, cultivate the health data ecosystem, and foster a culture of innovation.
- Establish mechanisms to better integrate HealthData.gov into broader HHS data efforts

2. Enhance the platform's capabilities to make the data more meaningful to a range of data customers and extend its reach

- Keep the website in a modern design so that users view it as fresh and well-maintained platform.
- Implement additional web features previously recommended by the National Committee on Health and Vital Statistics, including: providing recommendations for other data offerings that may be relevant based on other users' experiences and interests (“other users who used this also look at...”), enhancing the applicability of the content by allowing users to identify their roles (e.g. provider, employer, consumer, etc.), and providing users with drill down menus that display characteristics such as the format and metadata, allowing for customer ratings of datasets.
- Improve the automated processes currently in use to extract information from the open data portals that house the data covered in HealthData.gov's directory to ensure that HealthData.gov entries are timely and accurate.

2. Enhance the platform's capabilities to make the data more meaningful to a range of data customers and extend its reach

- Show relations between datasets to better represent datasets in a family; for example, datasets with multiple years should be included in one place.
- Publish corresponding data models, to help users maximize the information.
- Make data documentation consistent across HHS datasets, consistent with international standards and following the National Committee on Health and Vital Statistics' past recommendations.
- Develop a toolkit to enable communities to synthesize and integrate multiple datasets to address local health challenges consistent with the Office of the National Coordinator for Health IT's national interoperability roadmap, thereby being a national resource and engine for community and national learning health systems.

3. Develop a long-term HealthData.gov infrastructure that allows it to tap into state-of-the-art knowledge about big data, machine learning, and algorithms for increasing customer use

- Provide a foundational infrastructure to enable long-term planning and use of modern technologies to support the mission of HealthData.gov.
- Ensure that HHS has the technical capacity to use machine learning, data analytics, information and communication technologies, and other state-of-the-art scientific techniques to support users' needs.
- Establish a governance process to establish standards for the datasets that are referenced and other aspects of HealthData.gov.
- Develop business practices to allow operating divisions to keep the data comprehensive and timely, better understand and support user needs, and meet cyber security and privacy regulations.

3. Develop a long-term HealthData.gov infrastructure that allows it to tap into state-of-the-art knowledge about big data, machine learning, and algorithms for increasing customer use

- Create a Chief Data Officer who can oversee HealthData.gov activities and routinely engage staff within HHS operating divisions to solicit their input, provide feedback on how data have been used, and have a sustained commitment to releasing public data.
- Redefine HHS workforce roles to enable more active data customer support and foster health data community growth and a learning health system.
- Engage HHS in long-term data standardization activities to increase the ability to link datasets across operating divisions and decrease duplication.

4. Implement mechanisms to solicit data customers' input regularly to optimize their use of HealthData.gov and development of data-driven health innovations

- Create a “data product team” to engage regularly with diverse data customers from the private, public, academic, and other sectors to solicit their input on how to improve usability, access, and fitness for use.
- Solicit stakeholders' input systematically through surveys and focus groups.
- Use web analytics to improve understanding of HealthData.gov customers and their use patterns, thereby facilitating better tailoring of content to users, evaluation of dissemination strategies, marketing of data to specific groups, increasing the findability of data, and other benefits that result from enhanced data for decision-making and customer engagement.

5. Implement data stewardship practices to keep the commitment to the public to protect health data security and confidentiality

- Assign responsibility for data stewardship.
- Design processes for monitoring data recombination and the risk of mosaic and other similar effects that would permit identification or stigmatization of individuals or small groups.
- Design processes for monitoring downstream data transfers and use to avoid re-identification risks or violation of other principles of data stewardship.