

Testimony for Little Hoover Commission public hearing reviewing the nature and quality of interactions between California state government and the public
May 28, 2015

What CHHS Open Data Has Accomplished

The California Health and Human Services Agency is increasing access to public health, health care, human services and other data, driving improved transparency and data-driven innovation that enhance government services at the state and local level, across the health and human services landscape.

The CHHS Open Data Portal was implemented for open access to valuable public CHHS data assets, while protecting confidential private data—which will never be available on the portal.

Developing the portal.

CHHS first developed the CHHS Open Data Handbook, modeled after the work of the New York Health Department. The Handbook clearly lays out processes, procedures, and definitions for the CHHS Open Data program.

The CHHS Open Data Handbook and its related artifacts are important tools to communicating the business value of making health and human services data more accessible to the public, while also providing confidence to internal and external stakeholders that private, client specific data remain safely behind secure state firewalls. The Handbook has been designed in a manner so that it can be adopted by additional departments, both within CHHS and by other governmental entities.

CHHS is using a phased approach to implement its open data program. The CHHS Open Data Portal is presently in a pilot phase focused on the addition of CHHS departments in a sequenced schedule to begin publishing open data. CHHS is comprised of 12 state departments. The California Department of Public Health (CDPH) was the first to pilot open data in summer of 2014, followed by the Office of Statewide Health Planning and Development (OSHPD) in the fall of 2014. Two more CHHS departments (Department of Health Care Services and Department of Rehabilitation) have since on boarded to the portal and the balance of CHHS departments are scheduled, one per month, to begin publishing open data this year. There are currently 74 data tables and 57 visualizations available on the portal. The visualizations help inform visitors to the portal about potential uses of each dataset by illustrating how the data can be displayed in visual forms, including charts and maps. All 12 CHHS departments are expected to be publishing open data by early 2016.

The CHHS Open Data Portal technology employs a flexible hosting and deployment model. The primary value-add features of this solution is the application programming interfaces (API) that allow application developers to automatically consume the most current CHHS data into their respective applications. A comprehensive metadata platform also provides rich documentation describing

Testimony for Little Hoover Commission public hearing reviewing the nature and quality of interactions between California state government and the public
May 28, 2015

each dataset on the portal. Metadata is the critical component to open data that delivers meaning to consumers of these data.

Technical administration and organizational governance of the portal has been a shared effort amongst CHHS departments, with leadership provided by the Agency. This includes time commitments from a number of different skilled roles: data stewards, data scientists, program subject matter experts, technologists and project managers. Engagement of all of these staff is necessary to prepare and analyze the data to ensure we are posting quality data that will be useful to the public. This model of a CHHS “community of support” has served to engage staff and to make the most efficient use of limited staff resources, while still producing a high-quality product that has received positive press and interest from CHHS stakeholders and the media.

Data shared on the portal has already been publicized by major media outlets reacting to newsworthy public health matters. For example, The New York Times and KCRA-3 Sacramento, the local NBC affiliate, ran stories on measles immunization rates for schools across California using data from the CHHS Open Data Portal.

Operational improvements and culture change.

The CHHS open data initiative has laid the groundwork for important organizational improvements to enhance the performance of CHHS programs. While organizational resistance to change is natural and expected, open data has been a catalyst for CHHS programs to better understand the value of data and optimize data services to meet broader public needs.

CHHS departments have begun to develop new policies and procedures to advance data quality assurance and more timely data publishing.

Instilling the principles and benefits of improved data access, use and innovation has been an important step to gaining buy-in from staff, and in overcoming organizational inertia and resistance. When program leadership and staff see the value of open data relevant to their own work and goals, they are often eager to embrace open data publishing.

The experience of implementing an open data program in CHHS departments is generating new governmental partnerships that foster a more collaborative and dynamic workforce amongst traditionally isolated organizational units. This includes diverse functional areas, such as governmental programs, public affairs, data management teams, and information technology organizations.

Testimony for Little Hoover Commission public hearing reviewing the nature and quality of interactions between California state government and the public
May 28, 2015

The CHHS open data team facilitates regular interactions with data stewards and users—both internal and external. This creates opportunities for ongoing improvements to CHHS programs, and empowers program staff to innovate and to use data more proactively, and strategically, within their respective lines of work.

External partnerships.

CHHS has leveraged external partnerships to foster public-private collaborations that serve as a support ecosystem around open health and human services data. The ability to leverage these partnerships, community networks, and current talents have been critical to the early success of the CHHS open data initiative. Local community organizations help CHHS identify the value of data based on the needs of data stakeholders, including civic developers, data consumers, health data journalists, local governments, communities of practice, universities, technology startups, advocacy groups, philanthropies, nonprofits and others. In turn, these various data stakeholders focus on applying data to build useful tools and services, spurring innovation among open data partners.

In February 2015, a stakeholder representing a California non-profit firm (Civic Knowledge) contacted the CHHS open data team. Civic Knowledge is an open data research firm that helps governments and non-profits make better use of their data. Civic Knowledge, in partnership with the California Healthcare Foundation, is building a web tool that aggregates public data to tell visual stories and to provide interactive and engaging reports. The goal of this effort is to help county health policymakers better leverage data to make decisions. Because designations are not consistent across datasets (for instance facilities may have different names or number designations in different datasets), county policymakers often find it challenging to access data from various and disparate public data sources, including CHHS programs.

To populate this online tool with public data, Civic Knowledge is developing an organized, relational database that links multiple datasets into a useable and structured format. Because departments may use different numbers and names in their datasets, based on program needs, it's often challenging for users to link facilities, for example, across different programs' datasets.

The CHHS open data team, with knowledge of underlying program data, developed a "cross-walk" tool that links the data from across CHHS datasets. The tool allows users of the data, which are also provided on the portal, to be able to identify and match facilities based on the information contained in the crosswalk. The tool was immediately leveraged by Civic Knowledge in their community-level work, but is also made available on the portal for anyone to use. Staff from various CHHS departments now use the same tool to integrate data from across departments,

Testimony for Little Hoover Commission public hearing reviewing the nature and quality of interactions between California state government and the public
May 28, 2015

reflecting an important capability supporting the work of CHHS. The investment in the open data initiative drove the community connections that created this process improvement.

The California Health Data Ambassador Project is another example of the great benefit that external partnerships can provide to connecting CHHS programs with the needs of the community. Funded through the California Healthcare Foundation's Free The Data Initiative, the California Health Data Project embraces a mission to build bridges to local communities with CHHS data. Ambassadors encourage California communities to use the volumes of health data collected by CHHS departments in new, tangible, and relevant ways. Working with local community members, Ambassadors are committed to helping locals build and deploy at least one web application, data visualization, or information product using state CHHS data that benefits the local community.

The Ambassadors are regularly gathering feedback from health leaders in three pilot communities: Los Angeles, Fresno, and Sacramento. Within each of these cities, they bring together stakeholders from across sectors—local government, non-profits and community organizations, civic developers and technology companies—to ensure that health data reaches the people that can best put it to good use. Feedback is also circulated back to the CHHS open data team to help identify strategic opportunities for collaboration going forward.

CHHS has also partnered with the Sacramento Code For America Brigade to conduct a “code-a-thon” challenge for local application developers to make use of CHHS data in new and innovative ways. One result from this challenge was the creation of WICit, a web-based application that automatically pulls data from the CHHS Open Data Portal to display Women Infants and Children (WIC) program data about service location and eligibility information in a modern, user-centric web experience. Though not yet widely adopted, its development exemplifies the possibilities for state-local community collaboration.

Code For Sacramento also worked closely with the CHHS open data team to build out the CHHS presence on GitHub, an online collaboration platform for application developers. The CHHS Open Data Handbook is now available on this platform for other government agencies to leverage and adapt, demonstrating the CHHS commitment to broad community engagement and support of open data.

What's Next for CHHS Open Data

As stated, CHHS intends to finish the initial onboarding of all 12 departments to the CHHS Open Data Portal by early 2016. This effort will establish open data programs across the entire CHHS Agency, with departments continuing to regularly publish new datasets. This will expand the variety and volume of data available on the

Testimony for Little Hoover Commission public hearing reviewing the nature and quality of interactions between California state government and the public
May 28, 2015

portal, as well as expand transparency and encourage the cultural mindset of thinking about data in new and innovative ways. By harnessing the principles and benefits of open data internally, CHHS can improve efficiency, productivity, job satisfaction, and creativity within its programs.

Though many researchers agree that it is too early to measure specific outcomes or definitive health benefits from open data, there are many opportunities created by the institutionalization of open data through CHHS programs. So far, research from epidemiologist Cheryl Wold has shown that government is using open data to target health prevention and intervention programs, addressing problems such as food borne illness, hospital acquired infections, and childhood obesity.

The promise of open data is to change the way CHHS utilizes its data assets to improve the performance of government programs and make critical information available to the public. This expanded access to public data allows for communities to use data in decision-making to solve problems that government does not typically anticipate, identify, or address. In the broadest sense, open data serves to improve government transparency and public trust with California communities.