

ABSTRACT ACCEPTED FOR POSTER PRESENTATION AT THE ANNUAL MEETING
OF THE ASSOCIATION FOR PUBLIC POLICY ANALYSIS AND MANAGEMENT

**Assessing the Policy, Management, and Technology Challenges of Releasing Open Health
Data in New York State**

Erika G. Martin PhD, MPH, Rockefeller Institute of Government and University at Albany
Grace M. Begany, University at Albany

November 2015, Miami, FL

<https://appam.confex.com/appam/2015/webprogram/Paper13746.html>

OBJECTIVE: President Obama's 2009 open government directive encouraged government agencies to promote transparency by harnessing technology to make information accessible to the public. Government agencies at all levels have subsequently released thousands of de-identified datasets meeting specific "open" criteria: public accessibility, availability in multiple formats, free of charge, and unlimited use and distribution rights. Yet practitioners receive little guidance on how to release the data and face challenges of constrained resources and competing priorities. We aim to investigate the policy, management, and technology challenges that the New York State Department of Health (NYSDOH) has faced when releasing open health data. As a national leader on open health data release, New York's experiences provide insight into challenges likely to be faced by other government agencies and lessons learned.

METHODS: We conducted semi-structured qualitative key informant interviews with approximately 35 policymakers and practitioners at the NYSDOH, NYS Office of Information Technology Services, and Department of Health and Human Services to assess: (1) the policy, management, and technology challenges of releasing open health data; and (2) the perceived public value of releasing open health data to the public. Transcripts are analyzed using grounded theory, a systematic approach to discovering and reporting themes and concepts in qualitative data.

RESULTS: Challenges of releasing open data include: agency staff resource constraints, the complexity of the regulatory environment regarding personal health information (including multiple layers of federal and state regulations that may differ across datasets), required changes to internal NYSDOH agency culture (such as beliefs about data ownership and data sharing), extracting data from legacy systems, and data quality. Despite challenges, many key informants believe that opening data is promising. Initial benefits have included improved data quality (through processes to enhance data documentation and users identifying errors) and promoting internal agency data sharing. Additional long-term benefits might be the ability to analyze cross-sector linked data, discovering new uses of data beyond their original purpose (especially via data "mash-ups"), generating profiles of community activity, and increased scientific knowledge. Factors that have contributed to New York's success in becoming a national leader include a strong commitment from leadership, efforts to break down silos within the agency, strict guidance on the data release process (governance), and transitioning the open data effort from a special project in the Commissioner's office to a permanent program area. Open data platforms are still in their infancy and future work includes improving query functions and usability. However, key informants are overall optimistic about sustained commitment among leadership to the open health data initiative and its future success.

CONCLUSIONS: Health departments are keystones in the open health data ecosystem because public health agencies are one of the largest producers of personal, organizational, and system-wide data. Understanding their challenges and capabilities to release usable data to researchers and beliefs about the value of open data is critical to making recommendations about how to sustain a data ecosystem. By becoming a model community of practice, the NYS experience could influence open data activities in other states.