

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Leveraging best practices in data governance: An organization-wide data inventory and mapping project to support a five year data strategy

Travers, J¹, Campitelli, C¹, Light, R¹, de Sa, E¹, Stabile, J¹, and Iron, K¹

¹CPSO

Introduction

The professional regulation sector is moving toward risk-informed approaches that require high quality data. A key component of a corporate 2017 Data Strategy is the implementation of a data inventory and mapping project to catalogue, centralize, document and govern data assets that support regulatory decisions, programs and operations.

Objectives and Approach

In a data rich organization, the goals of the data inventory are to: enhance authoritative data that support programs; identify data duplications/gaps; identify data sources, owners and users; and, apply consistent data management and standards organizationally. Routinely used data assets outside the large enterprise workflow system (excel/word files; databases; paper collections) were catalogued. Using data governance principles and a facilitated questionnaire, departmental data stewards were interviewed about their generated data. Questions included data purpose/sources/types/formats/owners, retention rates, analytical products, gaps and visions for a desired data state. A data mapping methodology highlighted data set and variable connections within and across departments.

Results

To date, over 40 staff members in 10 departments were identified as data content experts. In addition to data in the corporate enterprise system, over 80 unique datasets were identified. In 1 large department, over 2,000 data elements across 26 datasets were inventoried. Data mapping analysis revealed thematic data domains, including member demographics, outcomes, certifications, tracking and financial data, collected and held in multiple formats ((Microsoft Access, Excel, Word), SPSS, PDF, e-mails and paper documents). While 72% of the data elements were formatted numerically, approximately 8% were free text. Significant data redundancies across staff members and departments were revealed, as well as unstandardized variable naming conventions. Gaps analysis highlighted need for standardized, electronic data, where not available and data management training.

Conclusion/Implications

Customized data mapping reports to data users will facilitate the development of local, standardized departmental data hubs that will centrally link to a centralized data repository to facilitate seamless organization-wide analytics, improvements in current data management practices and greater data collaboration with the ultimate goal of supporting risk-informed approaches.

