

A VA Executive's Guide to Open Data II

Office of Technology Strategies (TS), Architecture, Strategy & Design (ASD)
A VA Executive's Guide to IT Service Management



Introduction

In a past TS Note, we explored the principles, benefits, and challenges of open data and how the Department of Veteran Affairs (VA) shares data through the 2013 Open Data Initiative. The White House Office of Management and Budget (OMB) now requires government agencies to maintain data asset inventories. This TS Note will explore open data assets and their benefits to Veterans; the challenges and outcomes of acquiring, analyzing, and processing open data assets; and how you can contribute to the Open Data Initiative at the Office of Information and Technology (OI&T) at VA.

Overview

Data is a valuable national resource. The availability of data and the advancement of the computational resources that derive meaning from data have transformed the way business leaders and ordinary citizens make decisions on a daily basis. They are [data driven](#) and now *expect* probability predictions that were not available just a few years ago. The precision, accuracy, and timeliness that currently exist for determining weather patterns, network analytics, health prognoses, and customer service management have also elevated efficiency.

The Data Scientist and Business Intelligence

Not all data is valuable. Most of the data that organizations acquire needs [data cleaning](#) or scrubbing. An expert must ensure that corrupt, incomplete, inaccurate, or duplicated data is removed, to then process and present the data that *informs* – as *information*.

This is the job of the data scientist, who no longer sits in the back room! Indeed, the [Harvard Business Review](#) calls the role the “sexiest job of the 21st century.” “[Data scientists are the key to realizing the opportunities presented by big data. They bring structure to it, find compelling patterns in it, and advise executives on the implications for products, processes, and decisions. They find the story buried in the data and communicate it.](#)” This process is referred to as [Business Intelligence](#) or BI. BI uses technology tools to analyze data to determine actionable information to present to leaders - to help them make informed business decisions. The data scientist is a central figure who reports directly to leadership when delivering their most prized strategic asset – data!

Data as a Strategic Asset

Open data permits data scientists to obtain all the data they need as detectives; yet data can also appear to speak for itself. A classic example occurred at United Parcel Service (UPS), the world's largest delivery firm. In 2014, the data derived from a proprietary Global Positioning System (GPS) led officials to determine that the left turns made by drivers required more time, and were more dangerous when crossing intersections. UPS drivers are now required to make only right turns on new routes – routes that are also determined by data!

Data aggregation, the process in which information is gathered and expressed in a summary form for statistical analysis, delivers another form of data asset. It helps us determine the market value of real estate and the best prices available for airline reservations, to name a few. The data is aggregated into an application program interface (API) that is highly usable and programmed to query information.

The TS office within OI&T's Architecture, Strategy & Design (ASD) interacts not only with the ASD pillar offices, but also with multiple stakeholders within OI&T and with strategic offices across the enterprise. TS works closely with IT and business owners to capture business rules and provide technical guidance as it relates to Data Sharing across the enterprise, specifically for interagency operability.

For VA, data is a strategic asset, as a long-term means for achieving its mission. The development of data assets, such as the data that helps determine the Veteran experience, is the result of open data initiatives.

Open Data and Interoperability

Open data is data that anyone can freely access, use, and redistribute, with credit, or attribution, to whoever published it. Under the federal government Open Data Policy, agencies release valuable datasets that fuel collaboration across the public and private sectors. This transparency improves operational efficiency by creating deep analytical insights, measurements for policies, opportunities for the design of products and services, and interoperability. Interoperability enables systems to exchange and interpret shared data. It is important to VA, since VA is just one actor in a much larger eco-system. VA takes an active role in aligning with other organizations that also serve VA stakeholders.

A VA Executive's Guide to Open Data II

Office of Technology Strategies (TS), Architecture, Strategy & Design (ASD)

A VA Executive's Guide to IT Service Management

Metadata Standards

In order to make data available in searches, it must be accessible in a variety of readable and open formats. Open data must be descriptive enough to provide sufficient information for users to understand the data and its limitations. Thus, it must comply with standards for *metadata* - the data that is *about* data.

[Metadata](#) provides information that describes a data asset to make the asset easier to retrieve or locate. Metadata may include the purpose of the data, the time and date it was created, or the author.

The Future of Data at VA

VA manages a broad spectrum of data, with focus on improving data and data management. The plan for improvements includes:

- Managing waitlists at VA facilities to address long wait times
- Updating and maintaining Veteran records
- Sharing Veteran-centric data between the Department of Defense (DoD) and VA

What Can You Do to Get Involved?

The Office of Technology Strategies (TS), within the Architecture, Strategy, and Design (ASD) pillar at OI&T, is tasked to support the VA Open Data initiative.

To allow customers to have greater access to VA data and to maintain transparency, the Open Data initiative seeks new data assets. In order to comply with OMB requirements, the ASD Open Data team is required to maintain and continually add to both the VA data asset inventory and the API inventory.

VA employees can assist the Open Data team by submitting data assets that may be helpful and interesting to the public. Data assets can include any quantifiable information that is compiled within VA, such as a list or collection of information that is stored. An example of a data asset is the inventory of all Veterans buried in national cemeteries; this data asset is submitted every quarter by the National Cemetery Administration.

Non-Public and Restricted Public Data Assets

Data assets can be categorized by access among three levels: Public, Non-Public, or Restricted Public. Thus, it is not necessary for data asset submissions to be public facing. You can even submit metadata as a data asset, without submitting the data in the file that it describes. In these instances, the data asset would be classified as a [Non-Public data asset](#) in the VA's inventory. VA notifies the public that the data exists at VA, but that VA is not authorized to release it to the public.

Finally, you can submit a data asset that is partially released, or a [Restricted Public](#) data asset, because some information cannot be made public and is redacted, under the rules of the Freedom of Information Act (FOIA).

Data Asset Review

A Privacy Officer/FOIA representative will review your data asset submission to ensure that the data is suitable for public consumption and ready for release to the VA Open Data Inventory. The Open Data team updates the VA Open Data Inventory every quarter.

Working with the VA Open Data Team

In 2016, the VA Open Data team will reach out to the enterprise to discuss the VA Open Data Inventory, and to assist them in locating assets. If you have questions about data assets that may be applicable to this task, please contact Michael Miti-Kavuma (Michael.Miti-Kavuma@va.gov).

Conclusion

Developing data assets from open data is a process that assists organizations as a strategy to achieve outcomes. Data scientists clean, structure, and analyze the data to create valuable data assets to help make decisions and advise managers. At VA, data assets help to inform leaders how they can achieve their mission – to serve the most pressing needs of Veterans. With the availability of open data, business intelligence, interoperability, and your assistance; we look forward to using our open data assets to achieve this mission.