



Tech Insight III: Open Data

Volume 6, Issue 6

Office of Information and Technology (OIT)

In past *Tech Insights*, we explored the history, principles, benefits, and challenges of open data and how the Department of Veteran Affairs (VA) shares data through the 2013 [Open Data Initiative](#). We also laid out VA's plans to use open data to improve the quality of care for our nation's Veterans. This *Tech Insight* further explores open data assets and their benefits to Veterans; the emergence of new open data technologies; how open data could potentially be used within blockchain initiatives at VA; and the current work of the Open Data Initiative at VA's Office of Information and Technology (OIT).

Overview of Open Data

Organizations have long recognized that *data* stands out among their most valuable assets. In recent years, it is the *availability* of data and the advancement of the computational resources that derive meaning from data that have transformed the way Government and business leaders, and ordinary citizens, make daily decisions.

According to Brian Krzanich, CEO of Intel, “bits and bytes will replace petroleum as the primary fuel for the world’s economy.”¹ The “connected driving experience,” for example, makes drivers able to program their cars to take the most efficient routes, factoring for real-time traffic conditions. It allows drivers to seek conveniently located gas stations, restaurants, overnight accommodations, and parking spaces.

And what is the source of the data accessed for this fuel? Data access is strongly influenced by *open* data—the idea that *public* data should be easily accessible, freely available, and usable by anyone to develop products, and to republish as they desire—as opposed to closed data that is

¹ Source: Automobile Magazine, November 2017.

restricted to internal use by an organization (such as name, address, birthdate, etc.). Thus, open data is non-personal data; yet access to it has the power to revolutionize the way societies are governed and disrupt business commerce. Access to open data affords the ability to identify and predict large-scale trends and behaviors that provide new opportunities to gauge markets and improve services. Predictive analysis enables more efficient and effective healthcare preparedness and infrastructure planning.

Emerging technologies now use chunks of data as currency, most often known by Bitcoin, the proposal that made headlines when released in 2009. These technologies also track currency transactions made on peer-to-peer (P2P) networks, the communication overlay that transports the data. Security and robustness are their most important properties. Their analysis has been a powerful tool in the development of secure distributed payment mechanisms, secure storage, and the secure transfer and distribution of digital assets. In the United States and around the globe, this has led to the development of *secure* open data that is designed to share information and services between Government agencies.

A 2013 United States [Open Data Executive Order](#) mandated open data efforts across the health, energy, climate, education, finance, public safety, and global development sectors of the Federal Government. That same year, the White House also launched [Project Open Data](#) and several other open data initiatives that were designed to share best practices, examples, and software code to assist Federal agencies with opening data. In 2019, the [Open, Public, Electronic, and Necessary \(OPEN\) Government Data Act](#) was signed into law, providing a government-wide mandate for Federal agencies to publish all their information as open data—using standardized, non-proprietary formats. The OPEN Government Data Act built upon the [2013 Open Data Policy](#), *Managing Information as an Asset*, released by the Office of Management and Budget (OMB) and the Office of Science and Technology Policy (OSTP), making the key aspects of the policy permanent.

Open Data at VA

VA is a harvester of data. As indicated at [VA's open data web site](#), open data at VA is VA data that is freely available to the public; it is a by-product of the work that VA conducts while providing services to Veterans.

According to a 2014 [Projected Veteran Population](#) survey, the Veteran population in the United States can be described in exponential terms. From approximately 23 million in 2014 to 20 million in 2019, and a projection of just under 15 million in 2043, Veteran populations of this magnitude suggest that VA can be expected to continue to generate an abundance of data, spread across multiple information systems at VA, even in consideration of the possible decline. The data is quantified and categorized for statistical and inventory purposes. These categories include healthcare, education, physical asset tracking, gravesite location, and Veteran demographics. The greatest challenge for VA is that the data sets are vulnerable to threats, and if exposed, could be used in various exchanges within the VA network.

Benefits and Principles

As stated on the [Project Open Data](#) site, Government Open Data should be consistent with the following principles:

- *Public* – Agencies must adopt a presumption in favor of openness to the extent permitted by law, and subject to privacy, confidentiality, security, or other valid restrictions.
- *Accessible* – Open data are made available in machine-readable, convenient, modifiable, and open formats, that can be retrieved, downloaded, indexed, and searched.
- *Described* – Open data are described fully (through metadata) so that consumers of the data can understand their strengths, weaknesses, analytical limitations, security requirements, and how to process them.
- *Reusable* – Open data are made available under an open license that places no restrictions on their use.
- *Complete* – Open data are published in primary forms (i.e., as collected at the source), with the finest possible level of granularity that is practicable and permitted by law and other requirements.
- *Timely* – Open data are made available as quickly as necessary to preserve the value of the data.
- *Managed Post-Release* – A point of contact must be designated to assist with data use and to respond to complaints about adherence to these open data requirements.

There are many benefits to open data, well beyond those derived from routine consumer actions, such as the ability to book a flight using ticketing data from multiple airlines. From the Government's point of view, the business benefits of open data include the following:

- Increasing transparency and concomitantly, strengthening its relationship with the public it serves
- Saving time and expenses in responding to Freedom of Information Act (FOIA) requests
- Saving billions of dollars in operational costs by avoiding duplicative internal research
- Discovering complimentary datasets held by other agencies
- Empowering employees to be better informed by making data-driven decisions

From a global economic standpoint, open data can have even larger impacts. The increased efficiency it affords can result in a multitude of new products and services.

Submitting Data Assets

Data assets are 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 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. Restricted public data assets include some data that cannot be made public; that data is redacted under the rules of the Freedom of Information Act (FOIA). A privacy officer or FOIA representative reviews the 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.

Open Data and Blockchain

As of 2009, open data transitioned into a useable currency when the unknown person or group of people, known as Satoshi Nakamoto, devised a way to mine networks for blocks of data and created the first blockchain database. The new currency became known as Bitcoin and quickly took the digital trading world by storm. This was also the first known method for transferring data without it being copied for added security. Transactions were also made secure by using cryptographic methods of binding blocks of data together, with blocks of data carrying a time stamp of each transaction and identifying the parties involved.

Just like Bitcoin, [blockchain technologies are the next evolution](#) of the open data concept. A blockchain was designed to resist modification to ensure the integrity of the data. Simply put, blockchain is a public digital ledger that is used to record transactions across many computers so that any involved record cannot be altered retroactively, without the alteration of all subsequent records (called blocks). A record or block is a related group of data fields, such as the first name, last name, birthdate, and address of a Veteran. A blockchain is described as a value-exchange protocol used to exchange blocks of data. To assist in your understanding, you can think of the use of another type of protocol that you are likely familiar with, such as the file transfer protocol (FTP), a communication protocol that is used to transfer files.

Linking blocks of data together to form a chain is the premise behind this new evolution and the open data security that it provides. Blockchain relies on the consensus of a distributed group of computers along the chain. If one computer processes a block of data that does not carry the entire record or block, the transaction would be rejected. In blockchain, blocks are linked using cryptography. To change a block, a user would need to adjust all the previous blocks in the chain and gain acceptance from the group. Tampering with the blockchain demands notable computing power and control of 51 percent or more of all nodes, the points at which data intersects or branches in the blockchain.

Cryptography is also employed to secure the blockchain itself and secure communication between the elements of the system. It ensures participants can see the parts of the ledger that are relevant to them and that transactions are secure, authenticated, and verifiable.

Open Data and Blockchain at VA

While we know that blockchain includes blocks of open data bound together to form a digital chain that can be encrypted, and that it tracks transactions between users, what does this mean for VA and Veterans?

VA is exploring the use of blockchain to securely [exchange large volumes of data](#). This exchange has several applications which could impact the daily operations of the organization. Some of those applications include tracking and monitoring asset inventories and exchanging data from medical records, clinical trials, genomic data, and even mobile devices, wearables, and internet of things (IoT) connected devices. VA officials envision a future in which business operations can rely on blockchain and open data to track contract closeouts and have been testing the tracking of clinical appointments.

If blockchain data provides chunks of random data, how does an agency sort through it to find something useful? Open data must be descriptive enough to provide enough information for users to understand the data and its limitations, and it must comply with standards for metadata—the data about the 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, and the author.

VA manages a broad spectrum of data, with focus on improving data and data management. Using blockchain and other open data technologies is simplifying the way data is exchanged and how it is kept secure. For example, using blockchain in hospitals will encrypt data, track a patient's hospital visit, and help maintain health records and data integrity in one action, instead of requiring multiple systems and actions from users. Updating and maintaining Veteran records, managing waitlists at VA hospitals, processing Veteran claims for benefits, and sharing Veteran-centric data between the Department of Defense (DoD) and VA, are among the tasks expected to become more efficient.

VA's Open Data Initiative

OIT's Office of Architecture and Engineering Service is tasked to support the VA Open Data initiative, allowing customers to have greater access to VA data and to maintain transparency through seeking new data assets. To comply with the OMB requirements, the Open Data team is required to maintain and continually add to both the VA data asset inventory and the application programming interface (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.

VA shares its Open Data assets through its website, www.va.gov/data, where anybody can access data, APIs, tools, and resources that can be used to develop web and mobile applications, design data visualizations, and create stories directly from VA resources. The available data is a byproduct of the work VA does for Veterans (information such as VA facility

locations, homelessness resources, and family caregiver services) and VA ensures that no protected data is shared. VA encourages site visitors to use the data to create products, such as applications or infographics.

If you have data to share, VA's Open Data Team has been building relationships within the organization to identify data and add it to the [VA Open Data Portal](#). The Open Data Team would like to establish a relationship with your organization to increase the quality of datasets in the Enterprise Data Inventory, since you have access to a diverse range of data and information that would be mutually beneficial to publish. To learn more, you can contact the Open Data Support Team at OITOpenData@va.gov.

Conclusion

The use of open data enables organizations to develop valuable data assets. Information is now recognized as a national treasure; it is a strategic asset and resource to the Federal Government, its partners, and the public, and its online availability is mandated.

At VA, data assets help to inform leaders on how they can achieve the VA mission as they address the most pressing needs of Veterans. Blockchain is the next evolution of open data that will further secure processes across VA and other supporting agencies. With the availability of open data, blockchain, business intelligence, interoperability, and your data input, VA is better prepared to meet the needs of our nation's Veterans.

For more information on open data, you may reference the *Tech Insights* for [Open Data](#) and [Open Data II](#). For more information on blockchain, you may reference the *Tech Insight* for [Blockchain](#).

The *Tech Insight* Series

The monthly *Tech Insight* series aims to help readers make better decisions and be more informed customers of OIT products and services by providing them with high-level overviews of technologies that impact or will impact VA's IT environment. Each *Tech Insight* introduces topics in an easily digestible fashion by presenting background information on the topic, clearly explaining its importance within VA, and providing recommendations for success from OIT. [Access all Tech Insights](#).

Disclaimer: This document includes links to websites outside VA control and jurisdiction. VA is not responsible for the privacy practices or the content of non-VA websites. We encourage you to review the privacy policy or terms and conditions of those sites to fully understand what information is collected and how it is used.