G00248186

Moving Toward Data-Centric Government

Foundational **Refreshed:** 8 August 2014 | **Published:** 18 March 2013

Analyst(s): Andrea Di Maio

The confluence of open data and legacy modernization is creating the conditions for a new way for CIOs to structure government services and applications that is centered on data rather than processes.

Key Findings

- Open data is not synonymous with public data: An increasing number of organizations are leveraging open data for internal transparency and data integration purposes.
- Although the political capital of open data resides in its external consumption, government organizations can greatly benefit from it internally.
- New government digital strategies and transformation programs are giving data a much more prominent role than in previous years.

Recommendations

For government CIOs:

- Help establish a chief data officer role.
- Identify low- to moderate-risk legacy modernization initiatives where an open-data-centric approach can be used.
- Leverage some of the innovative open-public-data-based development methods for internal application development.

Analysis

Many government organizations are struggling with modernizing their legacy mission-critical systems by using a combination of commercial off-the-shelf (COTS), open-source, reusable and externally sourced (including cloud-sourced) solutions. At the same time, they are involved, to different extents and degrees of maturity, in a number of open government initiatives that often revolve around making public data more easily accessible through Web APIs.

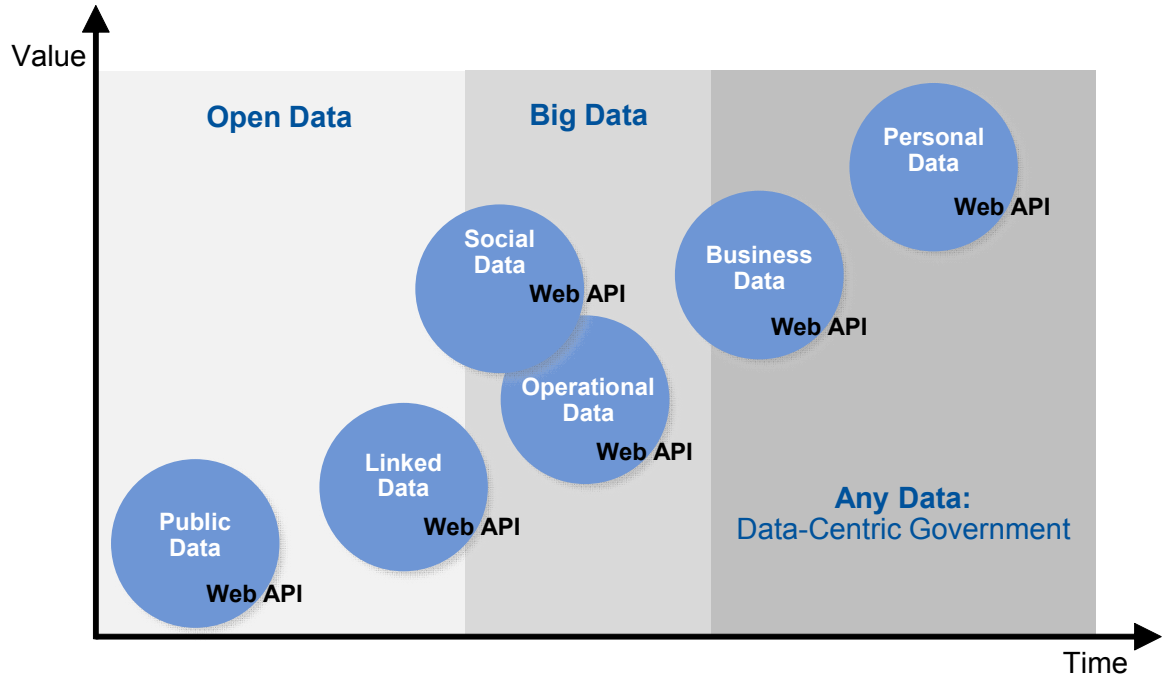
These two activity streams may be much more closely related than many people think. Better integration and data exchange between mission-specific applications and more-accurate data analysis capabilities to improve performances are top of mind for many government CIOs, and yet they rarely connect the dots between those problems and how open data may offer a solution.

Many tend to equate open data with public data, given its original definition (see Note 1). However data can be defined as open when it is machine-readable and is accessible through an API. This can apply to potentially any data that needs to be processed: whether it be public, discoverable through Freedom of Information Act requests, or restricted (for example, covered by privacy laws).

The value of public open data resides in increasing transparency, allowing internal and external parties to figure out new ways to use data that can deliver efficiency or even contribute to economic growth.

Figure 1 illustrates a potential path from open public data to open "any" data.

Figure 1. The Progression From Open Public Data to Open Any Data



Source: Gartner (March 2013)

Efforts are already under way in several jurisdictions, such as the U.K., to move from open public data to linked public data (see Note 2). The use of linked data supports easier sharing and integration of data across enterprise boundaries. Also, relentlessly opening data allows the uncovering of so-called "dark data" — that is, information assets that organizations collect, process and store in the course of regular business activity but generally fail to use for other purposes.

So far, this has been primarily a government concern. However, some enterprises (especially in the media, retail and telecommunications sectors) are starting to become providers of open and linked public data.

Besides government and business open and linked public data, people collect a wealth of data themselves through their own online communities. Gartner calls this data "social data," but it may well be open and linked in nature, depending on how those communities structure it.

Further, real-time operational data comes from devices that are embedded in industrial processes, vehicles and city infrastructures, as well as from consumer devices, such as mobile phones or consumer GPS, which add to open and social data and therefore help fuel the so-called big data phenomenon.

The ability to access this data as open data can unleash an even greater value than the simple use of government open data. There are already examples where mashing up data coming from consumer applications (such as Waze) or from social initiatives (for example, Neighborhood Watch) with open public data creates new ways of managing city infrastructure or increasing public safety.

The next, more disruptive step, is to consider usually restricted business-specific and personal data as open data. This data is not meant for public consumption, and is fraught with privacy and sensitivity issues, yet it can be modeled as open data to facilitate the development of more granular and agile applications, as well as a more coherent data exchange and analysis across agency boundaries and beyond.

There is only anecdotal evidence that this is happening (see the Evidence section), but this is a key trend for governments that want to become smarter — that is, more affordable and sustainable.

This approach gives rise to what we call *data-centric government*. The focus is no longer on applications, Data is now the key asset, around which applications are built.

Data centricity in government has several advantages. It supports:

- Better interoperability and joining-up. Rather than being forced to extract data from applications to achieve integration across organizational boundaries, data is described and accessible through a Web API by all prospective user applications according to specific access rights.
- Innovative application development and procurement. The same approach that is used to develop mobile or Web applications based on open public data (such as app contests, hackathons or datapaloozas) can be used to build applications that access nonpublic data. This would favor more-agile development and support greater employee centricity, as employees would be able to develop and/or compose applications to access data in more effective and convenient ways.
- The evolution toward "citizen data vaults," giving citizens a much better ability to control access to their own data and share it across agencies or with the private sector at their leisure (see Note 3).

Data centricity also carries a number of challenges:

- By helping to break silos, data centricity threatens the status quo, and thus may encounter all sorts of resistance, usually expressed in terms of security and privacy implications. Open data does not pose a threat to either per se, as access to data can be controlled in an even more granular and auditable way. However, the proprietary attitudes of programs or agencies that generate data can also have a chilling effect on open data and data-sharing initiatives. Data "ownership" issues are often as big a barrier as security or privacy, so open data governance is key.
- Open data and public data are often synonymous or used interchangeably, and this leads to possible confusion. Senior leaders do understand the importance of open public data and want to be seen as progressive and transparent, while they are unlikely to see any political capital in focusing open data on more-internal issues, such as productivity and integration (see Note 4).
- Required skills for data architecture are in short supply, and open data expertise is mostly available in the Web 2.0 and open government circles where people are mostly concerned with transparency and cool apps.

Recommendations

Government CIOs should:

- Promote the establishment of a chief data officer role, making sure that this position carries responsibilities that go beyond traditional open public data provision, and set metrics to encourage actual transformation. While the details of the position and the reporting line will vary across organizations, it is important that the emphasis is as much internal as it is external, with a focus on how to open corporate data for better internal use and more-effective application development, rather than on supporting more traditional open government endeavors.
- Identify low- to moderate-risk legacy modernization initiatives in which an open-data-centric approach can be used.
- Leverage some of the innovative open-public-data-based development methods for internal application development, mostly focusing on applications for increased employee productivity, as well as to deliver better information and services to constituents.
- Start securing the skills needed to support data-centric development.

Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"CEO Advisory: Chief Data Officers Are Foresight, Not Fad"

"Government Open Services Are the Next Step for Government Open Data"

"Meeting the Information Needs of the Government Business Executive in 2023"

"How to Determine the Value of Open Government Data"

Evidence

The U.S. Digital Government Strategy and the Danish government report "Good Basic Data for Everyone — A Driver for Growth and Efficiency" provide hints of an extension of open data beyond the realm of public data (see http://blogs.gartner.com/andrea_dimaio/2012/10/22/why-government-should-care-less-about-open-data-and-more-about-data). In the Netherlands, data centricity has been supported through a focus on "basic registries," which are now converging toward an open data approach: At least one significant legacy modernization project, at the Dutch Prison Service, is using this method.

Initiatives like the Blue Button by the U.S. Department of Veterans Affairs and the Green Button by the U.S. Department of Energy are further examples of this shift toward open data in government.

Note 1 Definition of Open Data

The most widely accepted definition of open data originates from a workshop conducted in 2007, during which eight principles were stated (see www.opengovdata.org/home/8principles):

1. Complete: All public data is made available. Data is electronically stored information or recordings, including but not limited to documents, databases, transcripts and audio/visual recordings.
2. Primary: Data is published as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms.
3. Timely: Data is made available as quickly as necessary to preserve the value of the data.
4. Accessible: Data is available to the widest range of users for the widest range of purposes.
5. Machine-Processable: Data is reasonably structured to allow automated processing of it.
6. Nondiscriminatory Access: Data is available to anyone, with no requirement of registration.
7. Nonproprietary Formats: Data is available in a format over which no entity has exclusive control.
8. License-Free: Data is not subject to any copyright, patent, trademark or trade-secret regulation. Reasonable privacy, security and privilege restrictions may be allowed as governed by other statutes.

In order to extend the concept of open data to nonpublic data, it is sufficient to tweak the first and sixth principles by saying that "all data is made available to those who have the right to access."

Note 2 Linked Data

Gartner defines linked data as a Web-oriented set of technologies and methods that simplify the publishing, discovery, interoperability and reuse of data for the purpose of generating information-sharing network effects (see "Innovation Insight: Linked Data Drives Innovation Through Information-Sharing Network Effects"). "Linked data" is a data management and mathematical principle, which holds that any time two data points are used together it creates a weighted link. As more valid use cases for the link emerge, the link grows in weight. Linked data is the principle that

this "weight" is further increased by the use of datasets in a pervasive linking pattern, and that the use of the sets of data begin to alter business processes for the positive.

Note 3 Citizen Data Vault

Citizen data vaults are services that provide data subjects with the ability to access their data outside the context of a particular government transaction, allowing them much finer-grained control about when and how data can be accessed, and by whom, within the relevant legal framework that they are subject to. They need to interoperate with government, as well as with third-party systems that directly provide services to constituents.

Note 4 The Open Data Paradox

A CIO in a city in North America that is well-known as an early mover in the open data space told Gartner that, while most of the value from open data comes from internal transparency and the ability of departments to use each other's data, this aspect has limited political capital because political leaders prefer to focus on the external use of open data.

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp