



Connecting the Dots: Getting Maximum Value From Data

MARKET TRENDS REPORT



CONFLUENT

Introduction

Responsive, relevant, timely, insightful. Agencies are asking a lot of their data these days. Data is the most critical element in all decisions agencies make. It's such a priority that every federal agency must now have its own [Chief Data Officer](#), responsible for managing and improving the use of their agency's data as described in the [Federal Data Strategy](#). State governments are moving along the same trajectory, with more CIOs than ever focusing on treating data as a [strategic asset](#).

It's a big job and a big change for agencies, which have been dealing with disconnected data silos, legacy applications and practices, and under-resourced data operations for decades. Making that shift from data as a passive to active asset takes some work, but it pays off. The result is widely available, reusable data that can fuel real-time insights and application development, and personalize every citizen touchpoint.

Making data work for agencies requires a modern data platform that can handle all data — not only real-time data, but batch data, data in the cloud and data in on-premises data centers. With a single architecture approach, agencies can make sure their data is available and ready to serve all user communities and applications, and continuously adapt to new demands.

To learn more about how agencies can use event-driven streaming to improve data access, collaboration and analysis, GovLoop teamed with Confluent on this report. Confluent's enterprise event-streaming platform helps agencies unlock and repurpose their existing data with a universal data pipeline based on an elegant architecture that makes data an active asset available to any credentialed user. This report will discuss the data access and management challenges agencies are facing and how to best address them.

By The Numbers

5.5/10

The rank citizens give satisfaction with government services. Government was the bottom-performing industry in every country.

\$89.3 billion

the amount the federal government is spending on IT modernization.

80%

of government organizations are still at the initial or developing digital maturity stages.

85%

of federal users say the ability to draw insights from streaming data is the key to modern decision-making.

92%

of organizations are finding success with microservices.

Top 5

Using real-time event-driven architecture is a top five emerging technology for improving customer experience.

67%

of government organizations are pursuing digital transformation.

“New approaches to gathering, analyzing and using data are transforming the way federal agencies fulfill their missions and serve the nation.”

- Data Governance Playbook
(Federal Data Strategy)

Removing the Friction From Data-Sharing

The Challenge: Barriers to Data Access and Value

Data is the engine that drives progress throughout government. When data is up to date and accessible, it can be harnessed to improve responsiveness, insights and decision-making. But getting to the point where all data is discoverable and accessible to everyone who needs it can be difficult, often due to the way the data landscape has grown over the years.

For many agencies, the ultimate goal is to better serve citizens and government workers by using the most relevant data to provide the best possible response “in the moment.” This requires being able to access a variety of data sets from a wide range of environments at all points in time. That includes transactional and biographic data from legacy databases, streaming data from IoT devices and other sources, and web click data, to name a few. The result is citizens who are more satisfied with government and decision-makers who have more insights and who are better able to collaborate and share data for greater positive government impact.

One of the most common issues agencies face is dealing with the sheer number of data repositories that have developed over time. Data in these standalone repositories, disconnected from other data sources and applications, is often difficult or impossible to access.

When it is possible to access data from some of these silos, it can take time and money to negotiate access rights and then build the right connectors. And in cases where data is housed in older, proprietary systems, it can be difficult to access it without negatively impacting the systems the data resides in themselves.

The standard data lake approach, which essentially combines all data into one resource accessible by all parties, doesn't go far enough. Data lakes are passive by nature and better suited for finding data or specific records. Instead, agencies need to ensure that all data is discoverable and accessible to all parties in the moment it is needed. Quite often, citizens and government workers need a blend of historical and real-time data if they are to respond in a manner optimized for the moment.

The Solution: A Universal Data Pipeline

An alternative approach is what's known as event streaming. Unlike data lakes, which essentially are snapshots of data sets at a particular point in time, event streams combine data coming from various resources into a single stream — making it possible to process, store, analyze and act on data as it's generated in real time. This approach makes it easier to combine both historical and real-time data from multiple sources into any application, providing an easier path to analysis and productivity.

“The story over time tells you something that each snapshot itself won't tell you,” said Will LaForest, Public Sector CTO at Confluent, which offers an event streaming platform. “The idea is simple: Every time there is a change to a source database, that change is distributed to everyone who cares. In other words, it is actively pushing data to the people who need to solve problems instead of treating data as a passive asset.”

This approach decouples the data producers from data consumers. Historically, users or developers who needed a specific data set would have to locate the source system and negotiate how to get the data. Because the data is now independent of its original owners, it is much easier to find and access.

Once a data source is connected to an event-streaming platform, that data is always available and up to date. From that point on, any application can access it via the event-streaming platform.

Now easier to find and access, the data can be used to greater advantage. For example, agencies could build a 360-degree contextualized picture of a single constituent applying for citizenship or health care benefits, guiding and responding based on who that person is and what they need.

An event-streaming platform also can improve both inter- and intra-agency collaboration. “If I'm the custodian of a data set, I simply publish it, and I don't have to worry about who wants it,” said Jason Schick, Confluent's General Manager for the U.S. Public Sector. “You remove the institutional friction of sharing data: negotiating with other organizations, building interface connections and worrying about compatibility.”

Best Practices in Event Streaming



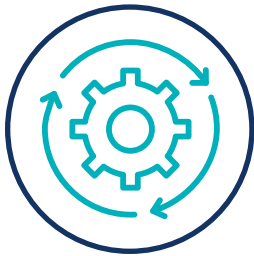
Consider an open approach to event streaming

Moving to an event streaming platform will go a long way toward maximizing the value of data today, but what about tomorrow? Inevitably, new data technologies will arrive with a need to access the data without creating new silos or compatibility problems. An event streaming platform, based on open source technology, will future-proof the enterprise because it makes the platform essentially technology-agnostic. Any application can subscribe and consume data that's published into an open source-based event-streaming platform based on Kafka. This makes it easier for agencies to share data, both today and in the future, which is necessary to execute on federated data-sharing efforts, where data is collected or exchanged across organizational boundaries.



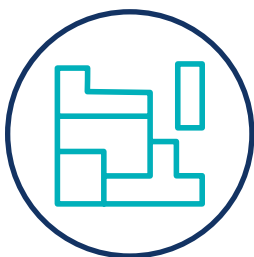
Change your data culture

While it's clear that data is increasingly important to government, harnessing the value of that data requires a change in mindset. A recent Gartner [report](#) found that culture and data literacy are the two top roadblocks for data and analytics leaders. In the report, Gartner stresses the importance of prioritizing cultural change and fostering a data-driven orientation. "You can make event streaming available, but unless people start thinking about data differently — especially about the importance of making data active rather than passive — they aren't going to make the best use of it," LaForest said.



Use event streaming to support DevOps

DevOps requires orchestrating across multiple services, resulting in a group of tightly coupled services. Typically, the more microservices in the group, the harder the service is to manage, and a break in one microservice could break the entire chain. With a more event-driven asynchronous model, microservices will decouple. That means that if one service fails, the service that it is calling won't fail.



Use event streaming as a building block for the future

While event streaming has many current benefits, those benefits will only increase over time. It is a key technology approach for anticipating problems and heading them off. Think about becoming better able to recognize the next pandemic and get ahead of it. "Once you have an event-driven architecture in place, you can be much more proactive," Schick said. "You can recognize things faster and put that data into motion to produce a lot of benefits for everything from safety and defense to improved voter registration and transportation."



Case Study

Flattening the Curve: Maximizing Value From COVID Testing

Last winter, when officials at the Center for Disease Control (CDC) began realizing that a new, potentially deadly virus was growing rapidly, they knew they had to get ahead of it as much as possible. Because the agency had worked successfully on many projects with Northrop Grumman for years, it turned to the systems integrator for help building a national infrastructure for all COVID line-level testing data as quickly as possible.

The goal was to create an authoritative source for comprehensive COVID testing data that would incorporate every test event happening anywhere in the United States or its territories. The system had to be able to deliver the data to downstream entities that were creating policy or engaging in rapid-response scenarios.

They used an event-driven architecture because it would allow the system to evolve and continue delivering value even as situations changed. The infrastructure, based on Confluent's event-streaming platform, combines multiple streams of data and creates a record for each testing event by a citizen. As records travel through the pipeline, they are enriched with metadata.

This process allows for a consolidated view of all situational awareness across streams. It also enables data to be aggregated and streamed to any destination at any time. The Northrop Grumman team could then organize automated "data storefronts," which help deliver data in real time to the White House Task Force and the Health and Human Services Department. A self-service data storefront also allows jurisdictions to view data.

HOW CONFLUENT HELPS

Confluent's Enterprise Event Streaming Platform helps agencies transition data from static to event-driven, from batch to real time, and from siloed to integrated. Built on Apache Kafka, the Confluent platform helps agencies modernize IT infrastructure, accelerate application development and connectivity, simplify operations at scale, boost citizen engagement, unlock the value of IoT, advance situational awareness, make SecOps more efficient and improve situational awareness.

With FIPS 140-2 compliance, comprehensive data encryption and role-based access control, Confluent's event-streaming platform has helped countless government organizations unlock and repurpose their data for modern applications and use cases.

For more information: <https://www.confluent.io/government/>

Conclusion

Report after report emphasizes the importance of data in agency decision-making, meeting mission goals and improving citizen services. Agencies must find a way to meet those goals while dealing with disparate stores of data, along with legacy systems and databases that are slow and expensive to access, and their existing data integration frameworks, which don't communicate well with one another.

Achieving all of these goals requires a new way of thinking and a new way of managing that data. Agencies can achieve all of these goals by embracing event-driven streaming, which provides a decoupled, real-time approach to data sharing and management. This approach allows agencies to better access, analyze and respond to all relevant data from past and present interactions in real time. The results are greater levels of responsiveness and agility, better and faster insights, and more effective inter-data sharing.



ABOUT CONFLUENT

Confluent is a Mountain View, California-based company founded by the original creators of Apache Kafka. We support a client base that spans all sectors of Federal Civilian Government Agencies, Defense, Intelligence, State and Local Governments and Education.

www.confluent.io
1-800-439-3207
info@confluent.io



ABOUT GOVLOOP

GovLoop's mission is to "connect government to improve government." We aim to inspire public-sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to connect and improve government.

For more information about this report, please reach out to info@govloop.com.



1152 15th St. NW Suite 800
Washington, DC 20005

P: (202) 407-7421 | F: (202) 407-7501

www.govloop.com
@GovLoop