WHITE PAPER

A Vision for Data-Driven Government

splunk> turn data into doing™

Executive Summary

There is no doubt that we have entered the data age, and this will impact government as much as the commercial industry. The amount of data that governments deal with and generate will increase significantly over the next few years, but this will present us with an opportunity to understand our citizens better and improve service and policy delivery.

To do this we need to take a data-driven approach to government by using data in new ways and overcome some of the inherent blockers to using data more effectively. Examples of blockers are; data silos, lack of skills, the lack of a data sharing culture and poor data quality. We also need to avoid the tendency to build new data infrastructure to overcome these problems, or to spend significant amounts of time in moving or replicating data, defining standards or preprocessing data.

A data-driven government is able to ask any question of its data, whatever the format, structure or location of the data. It would be possible to interrogate, link and consolidate data of all types to become more proactive and insightful. This 'data fabric' approach exploits current data repositories and sources that already exist and works in harmony with new, existing and legacy systems that generate or manipulate data. Fundamentally, it facilitates the linking of operational, service, initiative and policy data into a common data fabric as an overlay and not as a significant new infrastructure investment.

The more we understand our data and are curious and creative with that data, the better prepared we will be in government. A government that can 'ask any question of its data' will be ready for the data age and able to transform service delivery in these challenging times.

Introduction and Aim

Digitalisation impacts every aspect of our lives and every sector of our economies. It is driven by emerging technologies that are expected to play an increasing role in our lives. The Internet of Things, Blockchain, Artificial Intelligence and Edge Computing are data-rich technologies and will either generate a lot of data themselves or create the conditions for increased data production.

We live in a data age, and this is just the beginning. In a recent global survey (Data Age report 2020), Splunk found that executive and senior leaders expect the amount of data generated to be five times the amount it is today by 2025. Are we ready for this data age? Probably not.

During the COVID-19 crisis, governments have become acutely aware of the power of data. Real-time health data (Covid-19 cases, infection rates, etc.) are being routinely analyzed, and governments are adapting their policies based on this data. But this is only the beginning of the seismic shift of the data age. Governments now have an opportunity to evolve their services and improve the lives of citizens.

This paper aims to propose a vision for a holistic government data platform, driven by a shared real-time public sector data fabric that enables insight-driven policymaking for efficient public services and improved social value for citizens.

Current Approach

Governments across the world are waking up to the power and opportunity presented by widely accessible data. It presents an opportunity to improve outcomes for citizens, possibilities to make industry more productive and the potential to improve the efficiency of government.

Over the past few years, we have seen data strategies emerge around the world.[1] For example, the UK and Germany both released their own national data strategies recently, with dedicated sections on improving data usage in the public sector.[2]

1. Federal Data Strategy, United States (February 2020) https://strategy.data.gov/overview/
   EU Data Strategy (February 2020) https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy
2. UK National Data Strategy (September 2020) https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy
   Federal Data Strategy, Germany (January 2021) https://www.bundesregierung.de/breg-en/news/data-strategy-adopted-1845882

Despite this, in our recent Data Age report,[3] we found that only 23% of public sector respondents from multiple regions across the world are aware of the impending data volume increase, and 89% are not yet prepared for this rapid data growth. This means some governments will be left behind or miss this opportunity altogether.

It is clear that governments have not yet fully realized the potential of data, even if many of them understand at a high level that data is an important strategic asset. As a consequence, data is typically used reactively by governments rather than proactively to improve outcomes for citizens and reduce overload.

Governments also tend not to understand the potential value of the data they already hold. Likely, they are only focused on a fraction of data that is already highly visible, meaning they are missing considerable opportunities to improve services, learn more about their citizens and become more proactive and efficient. We refer to this as government 'Dark Data' meaning it is unknown, ignored and untapped.

Barriers to Governments' data usage

Why are governments not gaining value from their data? The problems are numerous, common and often well documented. A sample of some key themes are:

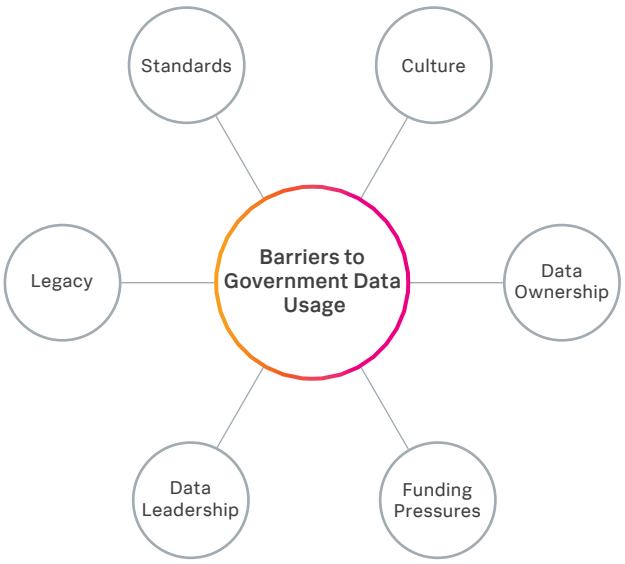

Figure 1: Government Data Challenges

- **Internal Culture:** Typically, there is no consistent data-sharing mindset in the public sector. Data is held in silos, and the default option is not to share the data across departments or functions. "Agencies collect data to serve their own needs, not cross-government uses."[4] Teams are risk-averse when it comes to data-sharing and a culture of fear of 'getting it wrong'. This is further exacerbated by a lack of consistent skills, policies, guidance and leadership. Using data is not front of mind in departments, and civil servants are not rewarded or motivated to use data in their work. This cultural obstacle is repeatedly highlighted in expert reports on data-driven government.[5]
- **Lack of common standards:** There is generally "a lack of consistency in the standards and systems used across the government, making it hard to share data efficiently", leading to "inconsistencies in the way data is recorded between departments."[6] However, unbending adherence to rigid data structures causes emergent scaling and compatibility problems. Data that may be usefully 'tapped' may not be able to be accessed due to unnecessary rigid access control, unsuitable standards or legacy data formats.
- **Legacy systems and processes:** Many critical systems have developed iteratively over time, increasing complexity and stretching the reuse of systems for services they were never built for. Manual reporting processes can provide immediate mitigation, but their proliferation builds significant barriers to real-time visibility and operational efficiency. Legacy systems still play a role but can have significant capital and resource cost, which clearly impacts the quality of services provided to citizens.
- **Data leadership:** Data skills are required at all levels of organizations, but executive teams don't necessarily have the experience of using data and are often unable to lead by example due to time pressures. Most governments do not recognize data experts or value their skills and data is seen as a specialist vocation.

3. The Data Age is here. Are You Ready: https://www.splunk.com/en_us/campaigns/data-age.html
4. Bringing Data into the Heart of Digital Government (Singapore, August 2019)
5. OECD Digital Government Review of Sweden, May 2019
6. UK National Data Strategy, 2020

- **Funding pressures:** Government budgets are increasingly constrained, but the pressure on services is growing, along with a demand from citizens for more and better services. The general response is a proliferation of approaches and a multitude of tactical funding initiatives around data. This has the unintended consequence of limiting flexibility at a policy level and vastly reduces efficiency unless it is driven by a consistent strategic approach.
- **Data ownership:** The drive for rapid change and technical transformation generates a requirement for mixed delivery models such as the utilization of cloud services or partnership with other third parties. This means governments have additional barriers to data access and visibility. This trend is accelerating, driving a need to ensure data ownership is always retained by the government and is leveraged for citizen benefit in a way that builds and supports public trust.
- **Poor data quality:** The proliferation of data systems and duplication of data across government lead to inconsistent, outdated and poor data quality. Data quality has a significant impact not only on current operations but more significantly on the ability to understand and react effectively to changes and emerging crises.

Data-Driven Government

A new vision for Governments

The **'Data-Driven Government'** platform (see Figure 2) is a real-time 'window' into the government's current data, whatever the format, structure or location of the data. This platform can interrogate, link and consolidate data of all types to present the opportunity for governments to ask any questions of their data and to become more proactive and insightful.

Data-Driven Government Platform / Data Fabric

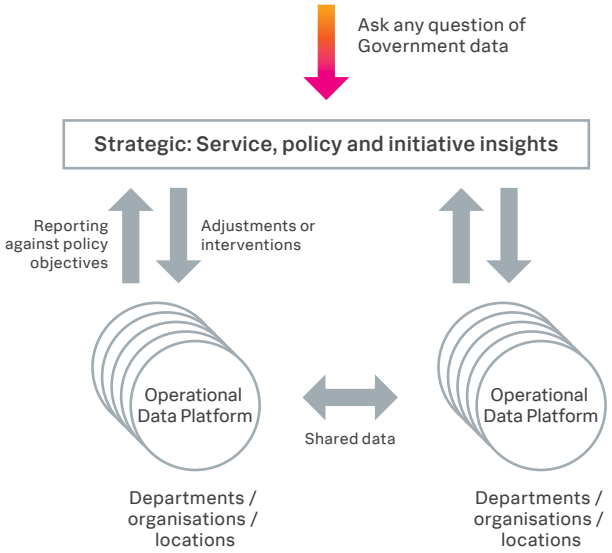

Figure 2: Data-Driven Government

An effective data fabric is built first from a focus on increasing the visibility and understanding of critical operational services. These **operational data platforms at departmental level** will provide a foundation for a real data fabric, generating immediate traction through tangible improvements and empowering internal front-line users. As greater operational efficiency is gained in each department, it can be expanded to discover new insights, consolidated into meaningful priorities thanks to continuous reporting against policy objectives (bottom-up approach). Data can be immediately shared across departments through the data fabric, resulting in increased productivity across government as a whole.

Key indicators can then be translated into **service and initiative insights**. These metrics can influence government policies and initiatives and the government can make adjustments or interventions (in funding, delivery, etc.) in real-time to meet the desired policy outcome (top-down approach).

By using simple queries, proven targeted machine learning and live data visualisations, it is possible to ask any question of government data and provide answers rapidly to internal and external stakeholders.

How to achieve this

Challenging conventional approaches

The data-driven government approach is holistic. It provides direct benefits to all levels (operations, service implementation, policy generation). It does not constrain governments, nor does it require massive amounts of effort in preprocessing, structuring, delivering new infrastructure or have a prerequisite for the creation of new IT programmes to deliver or consolidate 'data lakes' or 'data hubs'.

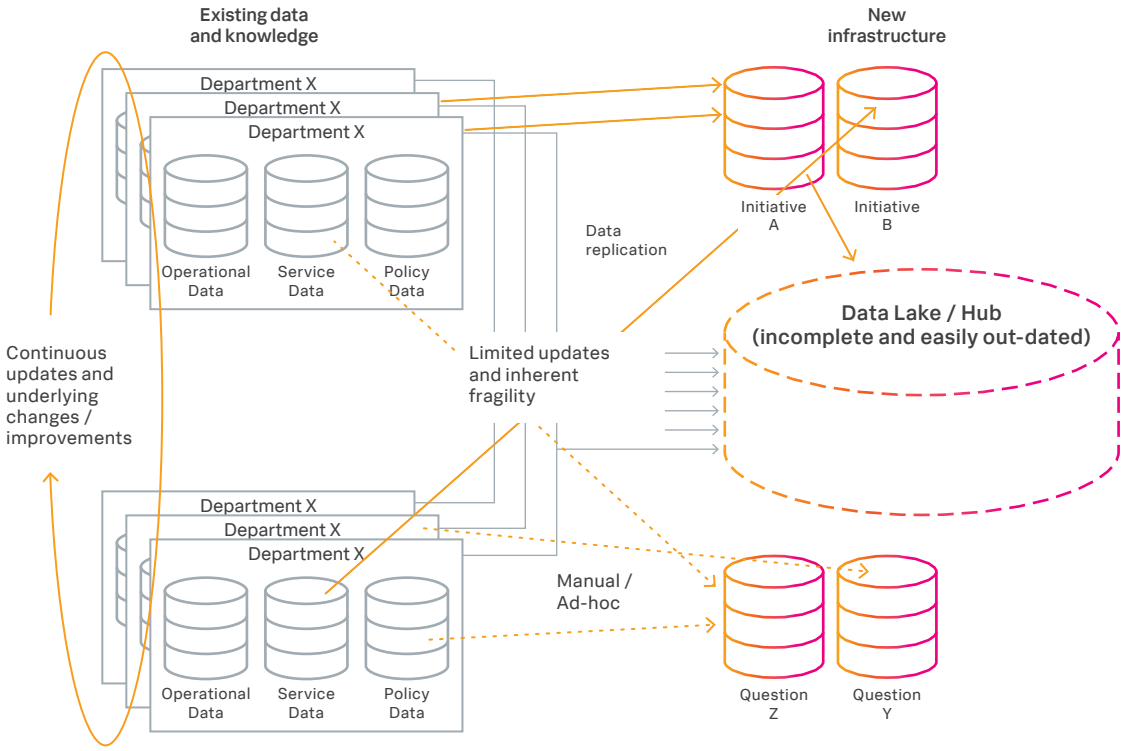

Figure 3: Conventional Data Architectures – Data lake / Hub

The disadvantage of these conventional architecture-led approaches (see Figure 3) is that they restrain flexibility, typically cannot easily integrate legacy data already stored in data silos, and require data to be moved and transformed en masse for every initiative and again for policy and/or insight requirements. Questions have to be answered in a manual or ad-hoc manner and initiatives or operations are vulnerable to problems with data replication or lack of updates.

A new approach

Instead the focus should be on the utilisation of data wherever and however it is presented and discovering rich sources of untapped data across the whole of government. This 'data fabric' approach exploits current data repositories and sources that already exist and works in harmony with new, existing and legacy systems that generate or manipulate data. Fundamentally, it facilitates the linking of operational, service, initiative and policy data into a common data fabric as an overlay and not as a significant new infrastructure investment.

This data fabric *overlay* enables policy delivery, outcomes and decisions to be driven by up-to-date consistent and consolidated views across all departments and regions, whilst recognising and supporting the different operational ways of working within departments.

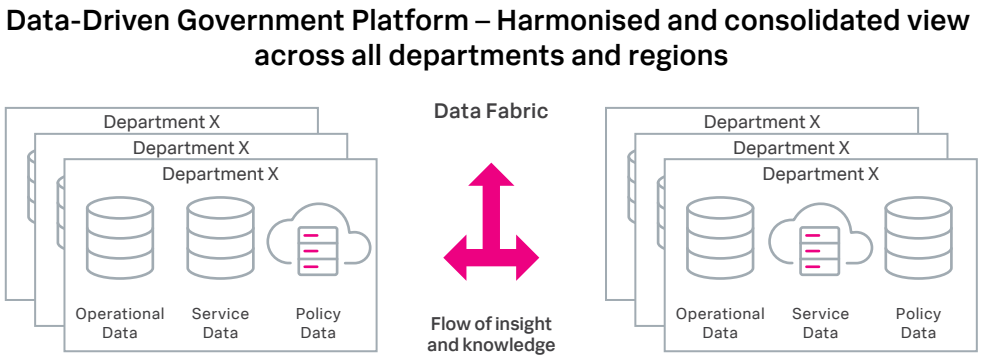

Figure 4: The data fabric / overlay

The underlying principle (see Figure 4) is to complement and connect relevant data, metrics and targets allowing the alignment of policy and interventions, without impacting critical operations or proliferating additional disconnected silos. It can be done in a transparent and auditable manner, ensuring the organization and the citizen have trust in the decision. Through this approach, the government can implement changes immediately and at pace without the vast upfront planning and investment in time and capital that other architecture-led approaches require.

With this approach, not everyone needs to be a data scientist, but staff should be encouraged and empowered to experiment with data, find new ways to use that data to make decisions and use data to better measure the impact of data-driven policy.

A framework is also necessary to encourage data sharing within public sector agencies and establish a data-sharing culture. To establish a data-sharing culture in the public sector, governments should consider developing a simple risk-based framework based on appropriate compliance and legal controls. Internal data-sharing should be the default option, and exceptions should be justified for significant national security interests or privacy concerns only.

Because governments typically hold and generate lots of data that they are not aware of, or are unable to take advantage of, teams could be given the tools to discover and tap into this data, to find new ways to improve decisions and citizen outcomes. Again, individuals and departments could be incentivized to do this, and metrics developed to track improvements.

Feature	Benefit
Real-time 'window' into data wherever it is located in government	A 'window' into the data whatever the format, structure or location of that data
Ask any question of the data	Become more proactive and insightful across government by using simple queries, machine learning and live data visualizations
Bottom-up, continuous reporting against policy objectives	Discover new data insights. Consolidate into meaningful priorities and easily report and analyse progress against outcomes
Top-down monitoring of service and initiative insights	Make adjustments or interventions (in funding, delivery, etc.) in real-time to meet the desired policy outcome
Operational data platforms sharing data across departments	Increased productivity and awareness both horizontally (department level) and vertically (strategically)

Figure 5: Features and Benefits

Benefits to Government

By being data-driven and harnessing all sources of data, the government will be able to observe the impacts of its policies and initiatives, orientate to ensure it has the best picture, adjust then deliver. However, there are other potential key benefits to this approach:

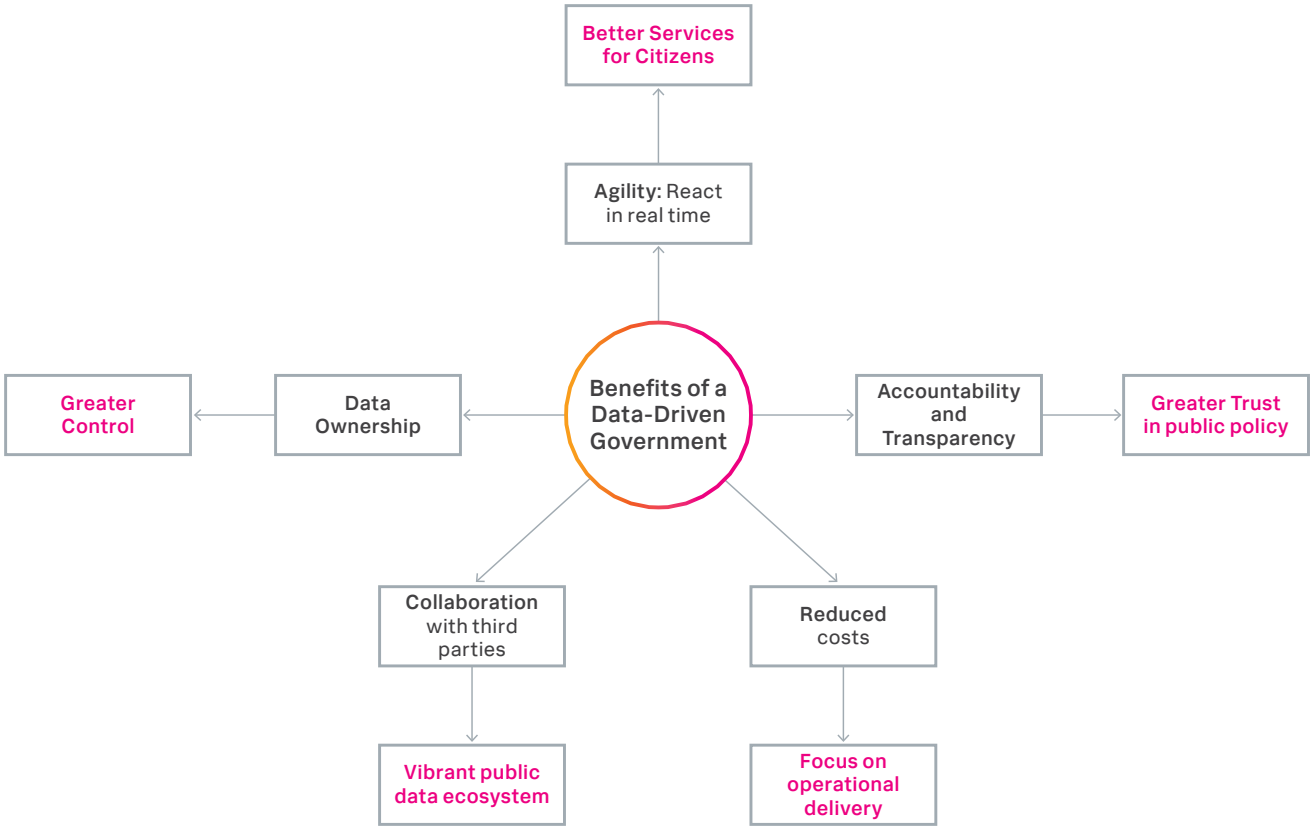

Figure 6: Benefits of a Data-Driven Approach

Improving the service to citizens

Governments can be more agile and proactive by being data-led. They will be able to analyze trends quicker, adapt policy and initiatives to react to better data insights and in effect, ask any question of their data.

Accountability and trust in public policy

With this approach, governments can demonstrate to citizens how data is used to set policy and deliver public benefit. Government can also experiment with policy and change course quickly. At a time when there is a level of distrust in government across the world, it is increasingly important to demonstrate that policy is not only driven by politics or ideology but also by the analysis of data. Citizens are more likely to trust institutions that manage to provide tailored quality public services.

Reduced costs

Governments don't necessarily need to lead with data infrastructure as their first investment or main focus. Data analytics can be an overlay to its existing infrastructure and services, not as yet another separate piece of infrastructure to develop and build. By doing so, it can save on infrastructure development and the requirement to move data across systems.

Data ownership

Data held by governments is a precious asset. This approach could allow governments to retain more control of their data.

Agility and innovation

This approach also allows for iterative initial adoption with selected departments or agencies as early adopters. It would provide the basis to innovate and experiment based on real data, without initiation overhead. As valuable insights and/or use cases are uncovered and proven, these can be shared and rapidly adopted by the rest of the government.

Collaboration with third parties

This data driven approach enables the government to provide answers to questions from external stakeholders or the public. For example, questions relating to citizens data trends could quickly be produced, sanitized and compliance-checked.

Conclusion

The more we understand our data and are curious and creative with that data, the better prepared we will be in government. If we can harness all our data, even the data that is dark to us, then we will no doubt improve our understanding of the needs of our citizens and improve our ability to meet these demands. A government that can 'ask any question of its data' will be ready for the data age and able to transform service delivery in these challenging times.

For more information, please contact pubsec_uk@splunk.com

21-17231-SPLK-A Vision for Data-Driven Government-A4-105-WP