



FROM 'CLOUD-FIRST' TO 'CLOUD-SMART'

How the Government can overcome the challenges and optimise the value of cloud-based analytics

Governments around the world have never been more focused on improving the quality and efficiency of their services or on driving up value for money. As a long-standing leader in data science and analytics, SAS has been working with government departments and others across the globe for decades to help turn data into insight and insight into better outcomes. Our approach is helping to transform the way governments work, improve the lives of citizens, as well as the security and economic prospects of nations.

The more efficiently analytics can be deployed, operated and managed, the greater the benefits it can deliver. Our experience tells us that a particularly potent way to achieve this is through the cloud. Over the past several years public cloud consumption has risen to dramatic prominence within the business sector. Its perceived cost efficiencies and scalability are now turning heads within government departments. However, before you deploy a cloud strategy for data and analytics, what are the critical considerations and challenges you should be aware of, and how can you maximise the real-world benefits for your department?

HMRC's Job Retention Scheme (JRS)

"HMRC adopted proven approaches and data sets to its risking work on the COVID schemes and used the existing SAS Viya and SAS Detection and Investigation services to improve user experience and enable HMRC to handle increasing volumes of digital transactions in managing risk. Getting this through did require a huge collaborative effort... we had about three to four weeks to get up and running."

In this paper we share insights on three key areas to help departments shift their analytics strategy from 'cloud-first to cloud-smart' to help improve services and deliver value for money through the better use of data and analytics on cloud. Let's kick off with a brief evaluation of where we are with cloud for government.

WEIGHING UP CLOUD

The mantra of 'cloud-first' for technology investments in government has been with us since 2013. The number of customer-facing and back-office services leveraging the advantages of cloud is increasing steadily with the exception of analytics, which is still rooted in an on-premise world. Why is this?



DATA SILOS

Analytics relies on a wide range of data sources that flow from myriad systems, often with different business owners.



USAGE PATTERNS

Several analytics teams may need to use the data simultaneously or in isolation which adds further complexity to any shift to cloud.



COMPETING PRIORITIES

Various stakeholders with differing priorities and ambitions for capitalising on the agility of cloud computing.



SECURITY

Concerns persist over cloud provider security, data sharing agreements and data classification in the cloud.

IS ON-PREMISE VIABLE?

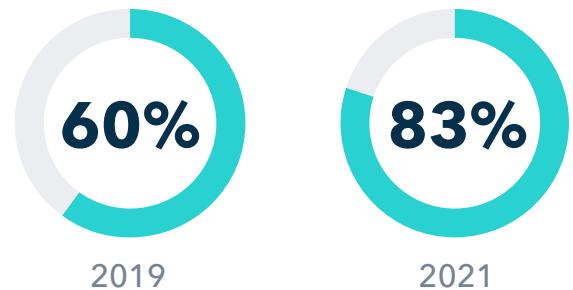
Continuing with the on-premise status quo also has its challenges. For one, as an increasing number of applications move to the cloud, bringing the data back on premise for analysis becomes a tiresome and expensive activity. This approach also creates an isolating effect. As government departments share more and more data to fight fraud and improve services, the agility that cloud analytics offers will be critical if multiple departments are to contribute to and benefit from a wealth of pan-government insights.

THE ROUTE TO SMART

With the on-premise model looking increasingly unappetising and cloud workloads on the rise it's important to be clear on your motivation. Not everything can or should go to the cloud. We advise carefully considering **why** you are moving to cloud when deciding **what** to move to cloud. A variation of the classic iron triangle of project management helps with thinking about "Why Cloud?".

For many organisations, the use of IT can either be to drive efficiency, doing the same thing while saving money (**Cost**), or innovation by speeding up value generation by working differently (**Speed**). Neville Cannon, research director at Gartner notes the different focus across government. "For example, national governments typically see cloud as a long-term pathway to strategic IT modernisation, whereas local and regional governments tend to pursue the immediate tactical benefits of innovation and cost savings". **Risk** is also important. You must get beyond the perception of risk to the facts and balance the risk versus the cost/speed benefits. We will explore this further below.

CLOUD IS EVERYWHERE



In 2019, 60% of workloads were running in the cloud and is expected to rise to over 83% by 2021.¹



From our experience of hundreds of cloud analytics deployments, there is no single answer for running analytics in the cloud, but the following topics - **Gravity**, **Security** and **Collaboration** - should serve as the key consideration waypoints on your journey.

1 GRAVITY

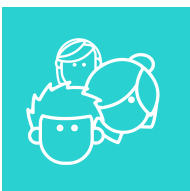
Data Flows & Owners

Analytics should be performed close to the data, something that is often called data gravity. However, data is often stored in a hybrid model, where data with a high security classification remains on premises, and other official data is stored in the cloud. Analytics is performed separately. In this scenario the challenge isn't the location of the data but the flow of data from one location to another. These flows - the data supply chain - drive costs both from cloud vendors (typically for transfers of data out of the cloud) and in terms of performance (network and compute) on large datasets and/or complex calculations. Cloud is not always the cheapest approach, so understanding the flow of data is critical in ensuring that transfer costs do not outweigh the saving on compute and storage costs. You will need to review your data wiring when moving analytics to the cloud since it will differ from your on-premise setup.



DATA LAKES ARE RARELY THE ANSWER

Understanding the data supply chain within your organisation (and its numerous systems) will help determine where analytics should be performed and on which data. We are seeing customers move away from data lakes after failing to see a return from Hadoop-based investments where, in the absence of specific data requirements for analytics, the approach was to store as much data as possible. In practice, understanding the strong data pathways is much more valuable to the business than having a big pool of data available.



KNOW YOUR DATA OWNERS

Another important factor in the cloud analytics journey is the engagement with data owners. It may seem obvious to say "identify who they are", but this can sometimes be difficult when analytics relies on so many different data sources. Typically, data owners' main concerns are around security, which we will address next. However, by engaging early to address concerns, understand a data owner's risk posture, and factor in specific security needs to policy and cloud architecture, you can help prevent delays or dissent as you move analytics to the cloud and achieve a successful cloud analytics outcome. Taking data owners through a structured risk assessment builds trust. National Cyber Security Centre (NCSC) is able to help and advise on this. SAS' Government [site](#) provides examples that can help provide confidence and strong precedence.

2 SECURITY

Federated Data & Shared Responsibility

Security is still the number one barrier for organisations moving analytics workloads to cloud. Concerns on general security, data loss and/or leakage are often cited as a reason to keep data on premises. In reality 60% fewer security incidents happen on public cloud compared to tradition data centres. Why? Reduced risk of human error through automation, consistently high-quality standards and frequent security updates.

Where data needs to remain in your own data centres it can be served up in the cloud for data science and analytics through a process of data federation. With federation the original data stays in-situ but users can access a virtualised and obfuscated view of it, according to predetermined security rules. We work with many organisations with sensitive data and there are hybrid cloud analytics design patterns that enable even the most secure organisation to exploit the agility of cloud.



SHARED RESPONSIBILITY

Security in the cloud is a shared responsibility. Cloud providers are responsible for security of the underlying infrastructure (compute, storage and network) and you are responsible for the applications and data that is run within the cloud. It is important to understand the shared security responsibility model to ensure you are protecting access routes to your data and not duplicating the underlying protection provided by the public cloud provider.

British Army invests in analytics leader SAS for digital transformation

“Technologies such as hyper-scale, multi-tenant cloud technologies, artificial intelligence and machine learning - and the ability to exploit the internet of things - provides the Army an opportunity to transform as part of the Fourth Industrial Revolution. We have already experienced how analytics can deliver significant insights and have delivered millions of pounds of efficiencies in the last two years.”

Major General JJ Cole OBE, CIO & Director of Information in the British Army



3 COLLABORATION

Repeatability & Governance

Cloud analytics opens up collaboration, not just for analytics and data science teams within the organisation, also across public sector organisations. Imagine if fraud actors detected by one organisation could be shared with wider public sector fraud teams to deter or prevent losses. However, data science is a relatively new discipline and many teams are struggling with the productivity challenge of getting models into production.

Analytics collaboration is much more than just sharing algorithms and models. It also involves bringing data analysts and data scientists together to share data and data pipelines across the complete analytical lifecycle. This spans everything from data wrangling through model development, model training and testing to production use, monitoring and re-training of models. The larger the community of data scientists and analysts the greater the potential for multiple ways of working with variable quality, auditability and re-use of models and processes.

The ability to spin up data science projects quickly and to try new tools and coding languages not supported by traditional on-premise setups makes cloud analytics ideal for experimentation. However, it is important to ensure that this speed and flexibility delivers the positive outcomes of collaboration and does not lead to wasted resources and effort. A key benefit from collaboration is the ability to re-use work that has already been done. This improves productivity and shortens time-to-value. Once best practice is defined and processes are supported by a proven governance framework, quality will be greatly improved through consistency. Ignoring governance could lead to the 'wild west' of analytics - a place no government department wants to visit.

“The inability to integrate analytic solutions into workflows and achieve frontline adoption is the number one inhibitor to why data and analytics initiatives fail.”

McKinsey & Company

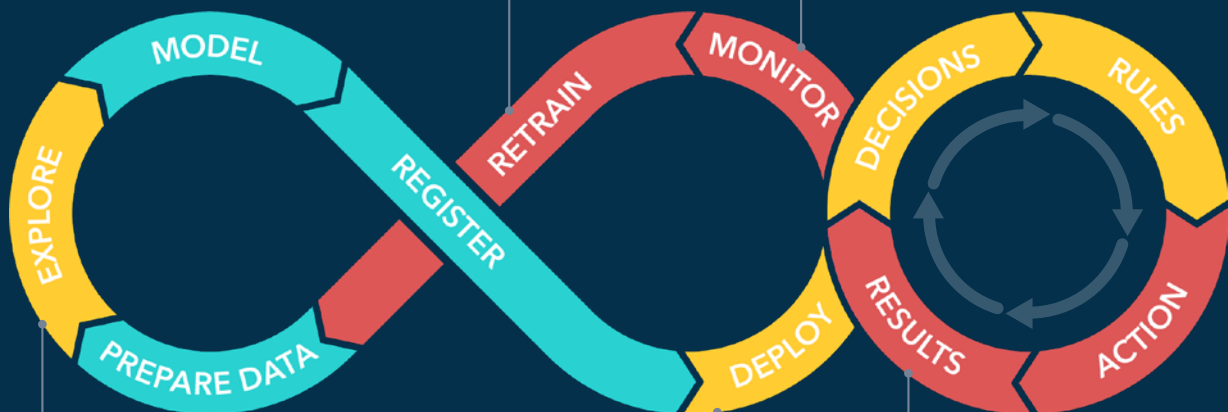
Addressing the challenges of collaboration: to accelerate time to value, improve quality and operational effectiveness of AI and analytics, leading organisations focus on the following:

- 1 Re-use** of data science assets created within the team and by others
- 2 Common tooling** and methods that will enable both data scientists and data analysts to collaborate efficiently. Putting in place common tools and methods to govern the analytics lifecycle is also a key building block to ensure data science efforts are converted into actionable insights that are trusted by users.
- 3 The last mile** of getting models into production is the hardest as the following illustration shows. The value of analytics is only realised when it gets into the hands of business users to improve core processes, up to that point it is cost without any benefit.

FROM DATA TO ACTIONABLE INSIGHT | COLLABORATIVE ANALYTICS LIFECYCLE

No central model repository for re-use. Little consistency & governance. Change of staff means loss of IP.

Very few model monitored/stress tested. Monitoring often a reactive exercise. Few models used champion/challenger method.



Some exploration but often time consuming.

Difficult to get open source models into production.

Tracking information about a decision difficult to obtain. Often very manual and inconsistent.

We are seeing pressure increasing for data science teams that have spent 12 months or more with little to show in terms of models in production. However, when the process works well, robust, auditable and collaborative production lines, can deliver exceptionally powerful outcomes. Two such examples include:



The audibility of deep learning models for chemotherapy response evaluation ensures transparent and explainable models leading to greater trust and adoption by clinicians.



A 90% reduction in data preparation time, 5x increase in model throughput and half the time to get models into production using our lifecycle approach and SAS tooling.

SHIFTING THE STRATEGY



We've seen that cloud analytics can deliver more value for government departments through greater agility and lower activation costs. The ability to collaborate more effectively across dispersed analytical teams drives reduced time to value for analytics insights. The costs of standing still can be high so departments should consider the following five points to help shift to a cloud-smart strategy:

1

Be clear why you are moving specific data and analytics workloads to cloud. If you have a speed/agility problem, cloud analytics is an ideal solution.

2

Understand your current and future cloud data flows before you start moving and consolidating data. Federation and data virtualisation techniques may mean you do not need to move all your data to cloud to realise benefits.

3

Identify your key data owners and walk them through a structured risk assessment, sharing stories of other similar workloads that have moved to cloud to provide confidence.

4

Architect your cloud security with care to ensure only the right people have access to the right data. Data should not be accessible directly, rather through a robust and secure analytics platform. Think about departments that share common data and architect data flows to avoid islands of duplication.

5

Put in place an analytics lifecycle management method supported by a proven tool. For example, SAS's ModelOps tools enable Open Source such as Python and R to be brought into a robust governance framework.

TAKE THE **NEXT STEP** WITH **SAS**[®]

If you are considering moving analytics workloads to the cloud, SAS offers a range of advisory services including Cloud Readiness and ModelOps Health Check. SAS also provides world class data and analytics solutions designed specifically for the cloud. We would be delighted to contribute to your thinking, please do contact us at our **SAS Government** site.

Philippe Vrijssen

Account Executive - BeLux Public

philippe.vrijssen@sas.com

Fabian.Ducheyne@sas.com

Account Executive - BeLux Public

fabian.ducheyne@sas.com