The Golden Rules of a Hybrid Analytics Strategy

Key factors every executive should consider and the pitfalls to avoid

By Peter Pugh-Jones, Head of Technology, SAS UKI

There's a passionate debate raging in the data science and user community about whether open source or an Independent Software Vendor (ISV) is best for analytics. We believe you don't need to choose – there is a third way that delivers the very best of both worlds. Let's discover how a hybrid approach to analytics can help you generate new business value and outmanoeuvre your competition while keeping you on track with governance. **How can you achieve it? Let's take a look.**

The headline makers of today's economy – the ones driving acute change at a macro level – seem to have one thing in common: they are capitalising on powerful technologies and trends to create platform-based business models capable of delivering unprecedented business value for themselves, their ecosystems and their customers.

So, as we all embrace this exciting new world (with a common understanding that advanced analytics is powering the insights upon which these organisations build their advantage), it's imperative that you evaluate what is the best analytics strategy for your organisation. And understand the golden rules that will help you successfully deliver it.

Before we do that, some context is important. Data science is proliferating across many of your lines of business. Yet data scientists are not a uniform bunch. They have a range of skills and undertake many different types of activity from managing data to developing machine learning models, deploying the analytics to delivering insights back into the business.

Each works with different types of data from multiple types of sources. What this is creating, right now, within your organisation is a hybrid of data science activities and processes. And this lends itself to a hybrid software environment – with many of your people wedded to either ISV or open source analytics tools and solutions.

From a business perspective, you must consider what links the disparate threads of activities, skills, software and data. What connects the projects being undertaken? What are the opportunities to drive value through repeating the use of analytical models? How is collaboration enabled? What about the project continuity and data governance?

This is the environment of many large organisations, and it can begin to deliver more value, faster when a unifying hybrid analytics strategy is at play. We'll discuss the golden rules you must follow to make hybrid analytics really work for you.

Stick or twist: identifying the winning hand in the game of analytics

Traditionally, analytics has been like a local football derby: do you support the reds (open source analytics) or the blues (ISV). Today, however, research shows that a very powerful hybrid alternative is rising in popularity. One that delivers the benefits of both worlds, while mitigating some of the risks and multiplying the benefits of open source, in particular.

The research paper in question, entitled 'Open Source Analytics versus Proprietary Analytics: What organisations need to know', undertaken by Vanson Bourne, notes that "most organisations are looking to expand their use of open source while retaining a solid foundation of proprietary technologies". This is because proprietary is optimised for operational and production analysis and includes integrated capabilities for data management, audit, governance and lineage plus much more, while open source can quickly bring new analytic algorithms to market.

Why Hybrid is the winning strategy

According to respondents, the use of open source is anticipated to bring new opportunities to both the **organisation (95%)** and to **customers (89%).**

While some **75%** would consider using an open source platform for data analytics, **95%** believe this poses several challenges for their organisation.

Striking the right balance: the key factors and pitfalls to consider before placing your bets on 'hybrid'

I believe that, as organisations of all kinds become more heavily dependent on advanced analytics, understanding the factors that could impact the success of your hybrid strategy is essential. Here are my top five rules for turning it into a valuable advantage.

Rule 1: Really know your business case

Although the general trend seems to be towards a hybrid approach to analytics development, assess what your business is going to be using analytics for. If you want to industrialise and scale the mission critical analytics you've already built in open source, hybrid will work well for you.
If you're embracing analytics-derived learning from artificial intelligence or machine learning, similarly you'll need a common understanding of the data involved and hybrid is ideal.
If, however, you're using analytics for bespoke open source-based modelling there is probably much less need for you to introduce proprietary software.

As a side, but nevertheless important note, make sure you're prepared to sell in the benefits of hybrid to those it will affect the most. The critical benefit is that hybrid gives them the ability to transition their ideas into largescale production environments turning their 'thinking' into competitive advantage much faster, rather than having to do so much of the heavy lifting i.e. coding.

This is beautifully demonstrated with a seemingly simple use case. Let's say your marketing team needs to analyse the reaction to a digital marketing campaign through sentiment analysis. You might write the programme in open source, but as your campaign rolls out across the globe you'll need to analyse reactions in different languages. With open source, you'd need to write a new analytical programme for every language. If you had a hybrid deployment instead, your people would be able to operationalise their original analytic 'capability' without having to reinvent the wheel with all the associated time and cost implications.

When scale and repeatability are critical factors and you need to cut time-to-insight, adding a foundation of enterprise analytics could give you the edge.

Rule 2: Understand the true costs, risks and mutual benefits

Vanson Bourne note that 36% of those asked, believe a combination of both ISV and open source solutions are believed to deliver the best ROI for organisations. By ensuring that you fully understand the core costs, risk and benefits you'll get the most from the hybrid approach.

The common perception is that open source costs less than proprietary software. Those surveyed in our research believe that open source can lead to cost savings of 57%. However, leaders must also consider the cost of redeploying skilled workers and recruiting others with different competencies.

Then there are the time and associated cost implications of working with open source analytics. There are many, many examples of organisations that have embraced open source without being able to operationalise it, resulting in huge financial waste and potential loss of competitive advantage.

Robin Way of Corios, an expert management analytics consultancy, notes many examples. One includes a top 10 insurance firm that chose to standardise all model development on one type of open source analytics software and all model deployment on another. After 18 months of work none of their model translations matched the results from their model testing. No one in the organisation knows how to fix the problem. This is far from the only example. In open source analytics environments, a lack of standardisation in working processes, methodologies and governance pose significant risks. When people leave the organisation, coding and project knowledge is therefore irretrievably lost.

Security and compliance can also pose risks. According to our research, 48% of people noted security threats as the most common potential vulnerability of open source analytics software. Non-compliance with new data regulations is also a potential pitfall to be avoided. If you do not have a common understanding about data sets, and cannot report where data has come from, your organisation could be at risk from new punitive measures. A good ISV system will include a data management framework that will address this issue.

Rule 3: Skill up, skill down, plan ahead for change

As analytic technologies and techniques evolve, it's crucial that you plan ahead both for cultural, organisational and skills change. Failure to do so could well slow down or even stall your progress, impacting your decisioning agility, customer-facing proposition development and overall competitive advantage.

Today, finding the skills to generate real value from analytics is proving problematic. Some 30% feel concerned about their over-reliance on a small number of experts and about what will happen should those experts leave. Anecdotal evidence from the Harvard Business Review puts this issue into sharp focus. Amongst a group of 150 specialist machine learning data scientists, only 50 had developed a functioning model, but none had managed to generate value from it.

At a practical level, if you are running a predominantly ISV analytics environment and adding in more open source analytics, you will typically require more programmers. If you are bringing more ISV software to your open source analytics environment, you'll need to recruit people with significant integration skills.

Rule 4: Build your strategy from data management up

As your use of analytics gains sophistication, you'll be working with new types of data. Think social media, geolocation tagging, IoT streams, voice logs and others.

Again, the Harvard Business Review puts much of the inability to drive value from data science down to the fact that "the data is a mess". Therefore, getting your data management strategy right from the start is imperative.

Ask yourself these questions to kick-start the conversation:

- Who owns the data?
- Where will it be stored/how will it be integrated into your existing infrastructure?
- What governance impacts will this have?
- What measures do we need to ensure data quality and security?
- How will we manage the increased internal demand for data?

You must consider how you will acquire the necessary data management skills: buy them in or work with a partner? If the latter, ensure they can help you bring together data silos, efficiently handle new data types and deliver a common understanding of your data to everyone, regardless of whether they are working in open source of proprietary languages.

Rule 5: Pick a partner with a solid track record and a compelling future vision

At SAS, we understand that open source versus proprietary is not a binary decision. With SAS® Viya™, our new, open, cloud-based analytics platform you can combine the power of SAS with open source languages and technologies. This allows your organisation to unify disparate tools, eradicate silos, increase productivity and drive agility. It's powerful and flexible enough to support all your analytics requirements from experimental to mission-critical.

And because it's a cloud-ready environment that serves everyone – from data scientists to business analysts, application developers to executives – you'll deliver reliable, scalable, secure analytics management and governance.

SAS® Viya™ drives value for your business by encouraging innovation through empowering staff to gather insights via the language that suits their skills. Teams can work together to solve complex challenges using a common and consistent interpretation of data, and multiple types of data from structured to unstructured, historical to real-time. They'll access those insights faster because SAS® Viya™ offers the repeatability and automation open source alone cannot and those analytical results will be sure and compliant because all assets and actions are fully traceable for audit purposes.

Let's sum up

The fact is, as analytics becomes the new battlefield for competitive advantage, taking a highly partisan approach will not maximise your advantage.

Evolving to a hybrid approach requires considerable evaluation comprising these factors.

- A clear understanding of the real costs, risks and mutual benefits of introducing or increasing open source or ISV analytics
- An appreciation of how your skills mix will change
- A willingness to sell in the benefits of hybrid analytics and plan for necessary process and organisational change
- The need to address data management issues head on

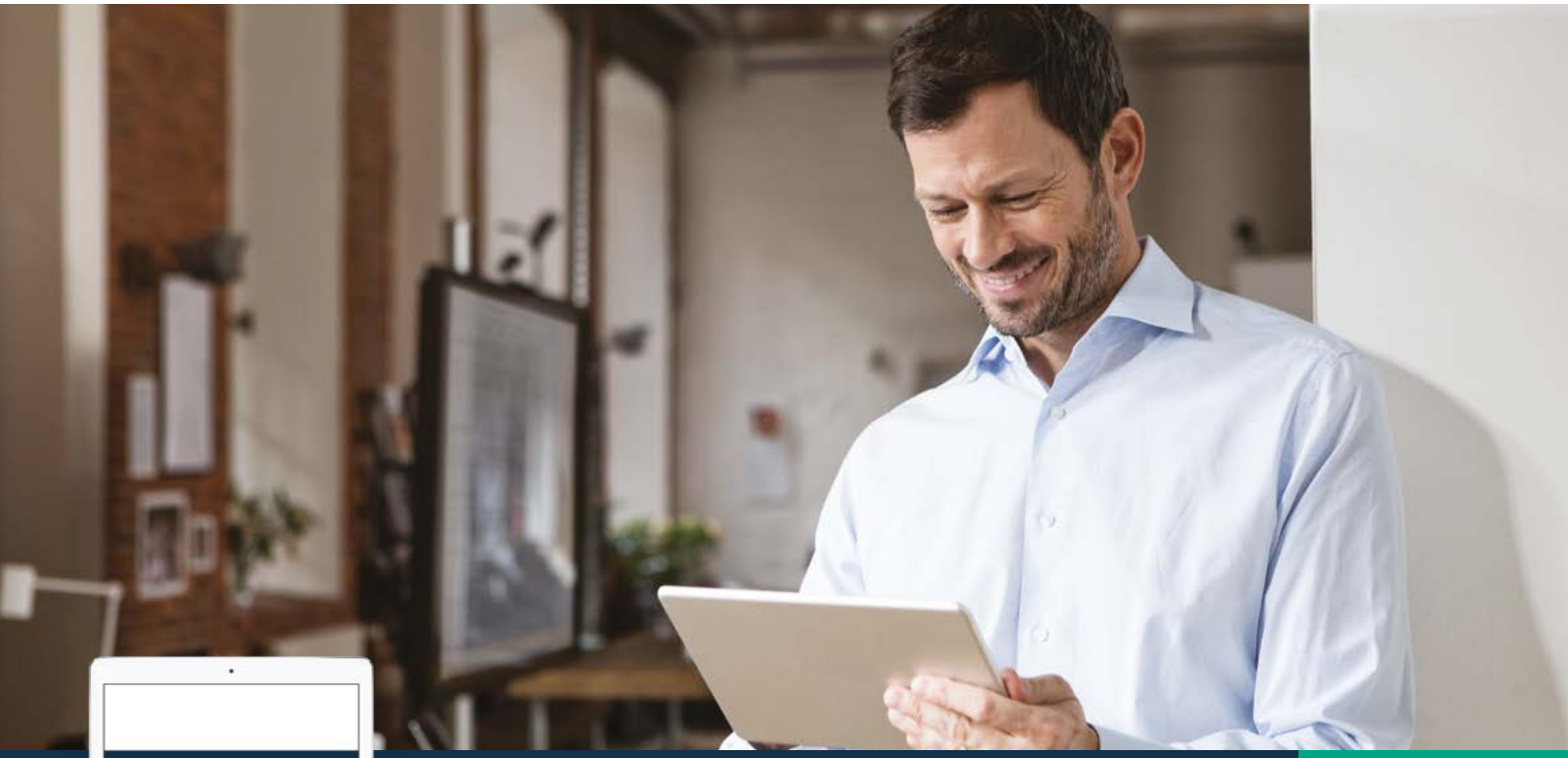
Follow these golden rules and you'll be heading in the right direction for greater hybrid success. The transition can be made much easier, achieved faster and with less business risk in partnership with SAS. We've been leading the field of productionised advanced analytics for a long time and it's this unique experience of supporting our customers in creating hybrid production analytics systems that continues to make us the leader in advanced analytics today.

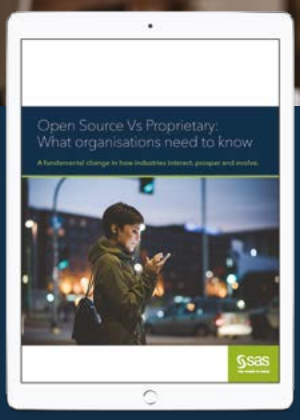
It makes sense, therefore, to embrace not only our technological advantage, but to allow us to help you address any skills and process gaps with highly experienced people who will work with you onsite to accelerate time-to-value and to map out your future needs.

Now is the ideal time to embrace the hybrid analytics future. It's a world where your data scientists have greater freedom to explore and uncover new competitive advantages.

It's one where speed to insight accelerates through improved collaboration. Where governance supports good business, rather than slows it down. Most importantly, it's one where you empower the data science function to deliver the kind of business value that has, until now, only lived as an aspiration.

We're ready to help.





Get the facts about the hybrid migration

Download our research Open Source versus Proprietary: What organisations need to know

sas
THE POWER TO KNOW.