



# Next- generation Insight Architecture

**Actioning Insights from Trusted Data**

June 2021

---

**KPMG.com.au**





Whilst improvements in available technology have lowered the barriers to entry for organisations looking to get better insights, the journey towards being insights driven remains elusive and challenging.

# Next-generation Insight Architecture

Organisations looking to drive value from data and insights have long invested in analytic assets – from on-premise data warehouses through to data lakes and machine learning platforms to name just a few.

While improvements in available technology have lowered the barriers to entry for organisations looking to get better insights, the journey to being insights driven remains challenging. Research suggests some of the factors are non-technical in nature and include things such as driving culture change and optimising the right operating model.

We believe some of these challenges can be alleviated by exploring next-generation insight architecture that is designed differently to contemporary approaches. Here, we explore the challenges and assess the potential of a 'data mesh' – a relatively new decentralised data architecture first described by Zhamak Dehghani, a principal consultant at ThoughtWorks.



## Single source of troubles

Building a single source of truth to enable business insights is conventional wisdom and a well-accepted goal for most data strategies.

After all, if there isn't a single source of truth then we risk duplicating effort, increasing costs and fuelling a general lack of trust in data and insights. In the early days of decision support systems, the approach of getting all data together in one place to analyse was feasible because the nature of insights generation was limited to management and standard business intelligence reporting. However, as analytical approaches have matured, building a single source of truth is increasingly problematic. Users of these systems have varying degrees of tolerance for truth and accuracy and other dimensions such as cost of solution and time to insights come into play. This presents a myriad of architecture challenges that are difficult to navigate while operating within the 'one platform to rule them all' construct.

This renders any one platform for the source of truth to be a challenging target state, because different users have different versions and tolerances for veracity of data.



## The challenges of monolithic platform architecture

Data architecture is about choosing trade-offs. No single insights platform is optimised for all use cases. While there are design patterns that can co-exist on a platform, experience tells us that a monolith platform-centric architecture is only ever optimised for one or two types of use cases. For instance, consider a data lake which is generally designed for ad-hoc data discovery and data science use cases.

Data lakes are optimised for fast-ish ingestion of data and making it available as quickly as possible for insight generation by data practitioners. It is however unfit for real-time integration or application of business rules to derive repeatable trusted management reporting. On the other hand, a data warehouse is useful for structured analysis of data, but cannot handle real time streaming events and complex event processing. A monolithic platform architecture regardless of whether it is a data lake or data warehouse (or even a hybrid data 'lakehouse') isn't setting organisations up for success.



## Business autonomy

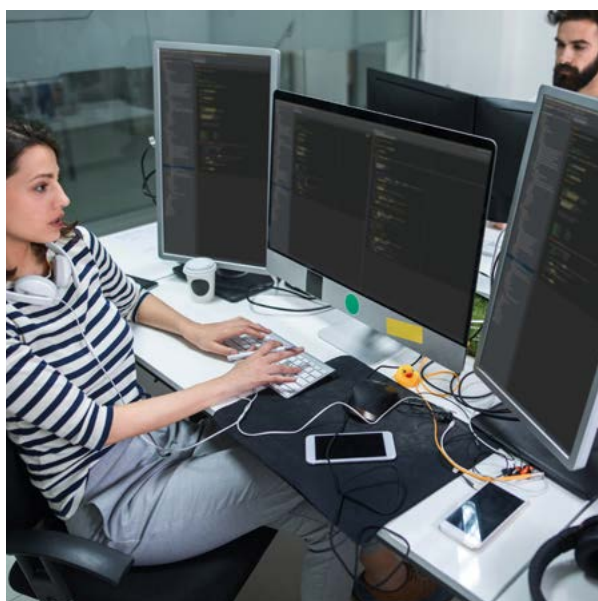
The traditional federated model where data and analytics teams aim to build a single source of truth, and business teams analyse this data, is being increasingly challenged as the available technology matures. This is caused by a few reasons. Firstly, the number of technologies that make up any enterprise architecture these days is increasing exponentially.

There causes a never-ending barrage of data ingestion needing to be done. Central teams tasked with building this are never really entirely up-to-date with the ingestion of the data, let alone delivering any insights or leveraging the power of machine learning. Secondly, available technology to analyse, visualise and crunch data is getting easier for average citizens to use and replicate. This results in business teams resorting to their own data manufacturing and refinement processes using a myriad of self-service tools because they can generate insights faster than it would take if they waited for the single source of truth to be made available.





A data mesh offers a distributed architecture that enables multiple teams to create data products, supported by self-service data infrastructure that can be provisioned at will.



## Data mesh: an emerging paradigm

Dehghani first published her idea of a data mesh in 2019. It explored the idea of constructing distributed, domain-driven data architecture supported by a product-centric development approach. We believe this idea represents the first differentiated approach to insights platforms in a manner that is designed to overcome traditional challenges. It is clear that Dehghani's software engineering background and experience in building microservices-centric application architectures have influenced her point of view on a data mesh.

## The promise – products not platforms

Centralising data and insights platforms is akin to building monolithic applications – they are slow, cumbersome and laden with architecture trade-offs that result in a suboptimal user experience for many data citizens looking to analyse, curate and shape data for a variety of use cases. A data mesh offers a distributed architecture that enables multiple teams to create data products, supported by self-service data infrastructure that can be provisioned at will. The aim is not for a landscape of fragmented silos of inaccessible and untrusted data – the data mesh approach requires product to focus on building domain specific data products that are provisioned independently and have distinct roadmaps.

This also enables different business teams to build products with differing governance requirements, latency needs and would enable them to make the trade-offs that come with speed vs. accuracy type decisions. Rather than ingesting data, Dehghani talks about serving data through a set of domain centric data products. The basis of her thesis is that building the same data lake/data warehouse type assets in the cloud will only bring the same problems we saw pre-cloud.

The architecture and the implied operating model is compelling. What's the catch and what can we do to leverage this today?

## Lingering challenges

In transitioning to a data mesh architecture, we foresee three main challenges that organisations will need to overcome.

### 01

**An organisation's ability to enable data serving instead of needing to ingest.**

Serving data across multiple operational and analytical systems requires a modern API fabric to be available. This is a challenge for most organisations that are not digital natives and have a significant portion of its technology architecture still running on legacy ERP or worse yet mainframes. This is further compounded by the fact that legacy operational systems often contain multiple data domains interspersed through the system. Uncoupling from these legacy data stores requires an API layer to be constructed around the legacy core applications.

As organisations modernise their core operations to run on cloud native applications, this challenge will be better managed.

### 02

**The ability to work across data domains.**

While it is possible to deconstruct data and insights solutions into products; discovering these products and analysing multiple domains across products is potentially challenging. For example, an organisation managing its supplier data domain is rarely ever just analysing the supplier domain. How do we analyse data that crosses multiple domains? Do we merge and create it into a new product? Do we centralise domains together to create, wait for it, a central data store?? Dahghani's architecture describes data lakes and data warehouses as just nodes on the mesh.

Therefore, if we adopt a product-centric model, do we still keep building these assets, or do we leverage existing assets through a modern API fabric?

### 03

**Organisational capacity and capability required to sustain multiple data domain-centric teams that build and manage products.**

This requires a very high level of data literacy across the organisation as well as a generally higher technical quotient to be able to manage domain specific data products over a long period of time. While it is getting easier for business teams to self-service their analysis needs, constructing a whole-of-business data domain-centric product structure is likely to be challenging.

Organisations need to consider how they can scale effectively across domain specific business teams.

The challenges associated with a data mesh are likely to reduce over time as technologies evolve and the general technology quotient in the workforce increases.

A data mesh lays the foundations of a new era, no longer just an aspirational goal, where cloud-based insights solutions can be designed in a sustainable, federated manner. We recommend organisations continue to explore how they can further move away from the traditional monolith architecture associated with data platforms and enable a more microservices based data solution, and product-centric development approach.

# Contact us

**Vivek Pradhan**

Partner, KPMG Digital Delta

T: +61 3 8663 8607

E: [vivekpr@kpmg.com.au](mailto:vivekpr@kpmg.com.au)

**Avi Sharabi**

Partner, KPMG Digital Delta

T: +61 2 9335 8929

E: [asharabi@kpmg.com.au](mailto:asharabi@kpmg.com.au)

**Danielle Malone**

Partner, KPMG Digital Delta

T: +61 2 9335 7850

E: [daniellem@kpmg.com.au](mailto:daniellem@kpmg.com.au)

**Lisa Jenkinson**

Director, KPMG Digital Delta

T: +61 7 3233 9577

E: [ljenkinson@kpmg.com.au](mailto:ljenkinson@kpmg.com.au)

**KPMG.com.au**

The information contained in this document is of a general nature and is not intended to address the objectives, financial situation or needs of any particular individual or entity. It is provided for information purposes only and does not constitute, nor should it be regarded in any manner whatsoever, as advice and is not intended to influence a person in making a decision, including, if applicable, in relation to any financial product or an interest in a financial product. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

To the extent permissible by law, KPMG and its associated entities shall not be liable for any errors, omissions, defects or misrepresentations in the information or for any loss or damage suffered by persons who use or rely on such information (including for reasons of negligence, negligent misstatement or otherwise).

©2021 KPMG, an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

Liability limited by a scheme approved under Professional Standards Legislation.

June 2021. 694658279MC