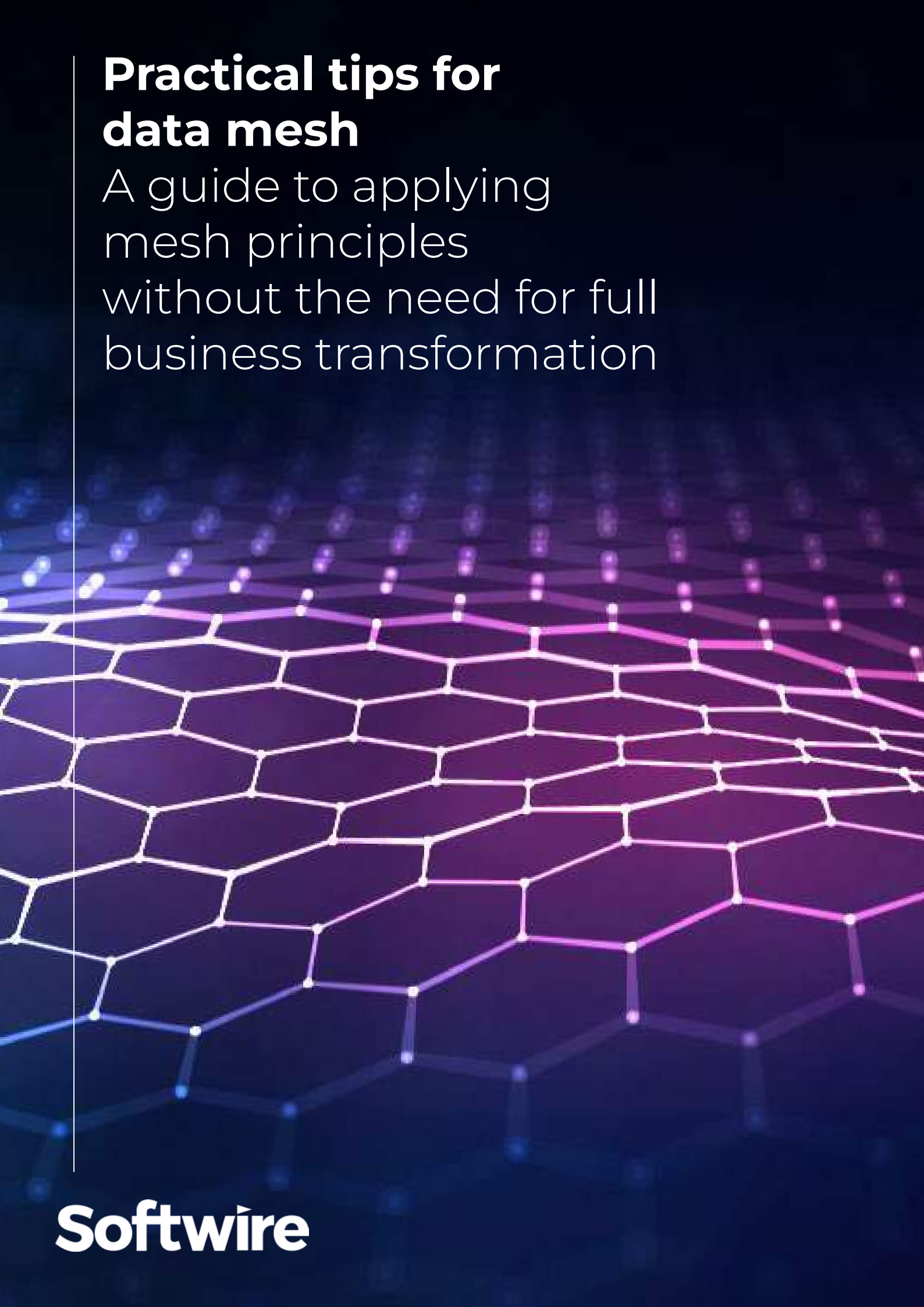
Practical tips for data mesh

A guide to applying mesh principles without the need for full business transformation



Softwire

Introduction

If you are a CDO, a data practitioner or simply someone who keeps up with the latest technology trends, you will likely have heard of data mesh.

While it might sound like the latest addition in a series of data architectures such as data lakes, lakehouses or data fabrics, data mesh is actually a modern approach to an organisation's data strategy that often requires a fundamental cultural change.

Data mesh applies ideas that have long been ubiquitous in software development (such as domain-driven design, building loosely coupled microservices and interaction through APIs) to the realm of data.

Who is data mesh for?

The concept of data mesh was introduced to achieve business goals that are eminently sensible: avoid dependency on a central, specialised data team; make it simpler and faster to share analytical insights across your business; and enable rapid implementation of ML/AI use cases.

Fully implementing data mesh can require large-scale restructuring of your organisation that will need long-term investment. If you aren't a large organisation struggling with the limitations of your current data strategy or are looking to expand your ML capabilities, you might think data mesh isn't for you and can be safely ignored.

However, we believe that adhering to data mesh principles can be of great use to any organisation, no matter how far you take them.

While the canonical (and highly recommended!) text on the topic, Data Mesh by Zhamak Dehgani, lays out a stepped approach for starting your data mesh journey to move away from traditional centralised systems such as lakes and warehouses, in the following, we want to give you some practical, useful takeaways that can be applied straight away – without requiring a full cultural transformation or giving up your existing data architecture solutions.

> “Data mesh principles can be of great use to any organisation, no matter how far you take them.

Overview of data mesh principles

The four principles behind data mesh — as laid out by the data mesh founder — can be summarised as follows:

1

Domain ownership: Business domains should consist of cross-functional teams responsible for both operational and analytical data, and hence data quality and data modelling, rather than outsourcing this to a centralised analytical data team.

2

Data as a product: Analytical data is shared as products: self-contained artifacts that contain everything that makes them usable (high quality) and valuable for a data consumer. Data products should have multiple output modes (ways of making their data accessible, for example via APIs, SQL or files) and can share data with other data products via an interconnected "mesh" made available through the self-serve platform.

3

Self-serve platform: To facilitate easy deployment and sharing of products across the organisation, a technical platform foundation needs to be built. This allows developers to deploy and data consumers to query and discover data products across the organisation. There is no prescriptive way or dedicated tooling (yet) for building such a platform – instead the focus should be on building the functionality that enables data mesh to function smoothly, for example through a series of distributed services.

4

Federated computational governance: Governance should not be a function completely divorced from the data itself, but instead include representatives from the domain, data providers and consumers, and the platform team. The platform itself should facilitate the automation of applying governance policies, such as access controls and GDPR concerns, to ensure data remains safe and compliant.

How can I apply these principles to my organisation?

Using data mesh principles will give you insights into how to improve current data processes, without requiring a full-scale cultural transformation. When we at Softwire are asked to help clients with their data strategy, we always start by asking the same questions:

1 Domain

- What areas of your business produce analytical data that is worth sharing? Is there analytical data that is already in high demand across multiple teams, with no clear owner?
- How well do your analytical and business domains (as defined in Domain Driven Design by Eric Evans) align? Have you defined the bounded contexts for your analytical data and do you have a good understanding of your analytical data models and their applicability?

2 Product

- Are there any problems with your data, for example quality? Are these restricted to specific data sets, or more widespread? Who could fix them?
- What are the most commonly used formats in which people in your organisation require data?

3 Platform

- What is taking up time in preparing and consuming analytical data, and which processes would you like to automate? Are there competing data demands that mean your request for more insight cannot be actioned?
- Is your current analytical data easily accessible, and discoverable across your organisation?

4 Governance

- Who at your organisation cares about your data?
- Who is responsible for access controls, and are existing procedures followed?

How to get value out of data mesh in any organisation

One of the biggest hurdles for people engaging with data mesh is the fact that a lot of organisations are still reliant on a central data team and central data store. Hence, they might not yet be ready for a cultural transformation to enable data products to be developed in a domain-aligned way, overseen by a federated governance function.

However, by asking yourselves the questions above, you might find that small changes to your strategy are possible that bring you closer to the data mesh goal of rapidly sharing analytical insights within a robust, governed process.

By using a data mesh mindset in thinking about data strategy, we propose a version of a "minimum viable data mesh": changes that your organisation can implement and validate quickly, which provide immediate value and that will provide a good starting point should you decide to go "full mesh" one day.

Minimum viable data mesh

1 Domain

You don't have to restructure your existing teams, but you should identify your analytical business domains, and map out how they fit into your wider business capabilities. Identify people who care about specific data sets and make them responsible for assessing the usefulness of any data products you introduce and ensure they feed back to the governance function.

2 Product

Identify what the most common use cases and formats for analytical data in your organisations are, especially ones that affect many different teams. Even if you only introduce a small number of shareable data products for your most common business needs, it could save you hundreds of hours down the line, as well as laying the foundation for the future.

3 Platform

Don't feel like because you are using technologies that go against data mesh "on paper" (warehouses, lakes) that you can't make use of the most important aspect of a new self-serve platform: making it easier to share analytical data across your organisation. Using your identified key data products as a guide, find tooling that solves the most time-intensive aspects of currently generating and finding analytical insights, even if this doesn't address all of the officially recommended aspects of a data mesh self-serve platform.

4 Governance

Naming governance as the last principle risks making it sound like the least important aspect of data mesh, when it is actually the most important for ensuring long-term success. Achieving automated policy and access management might be a long way off, but putting a cross-functional federated governance team in place will ensure that you keep critically evaluating your data strategy. Use regular governance meetings to discuss whether your data products meet your needs, what could be improved and if adopting more parts of data mesh best practice would be beneficial.

Getting started with data mesh

So how do you begin to apply these principles, and start building your first minimum viable data mesh? Guided by the questions above, you should find that you can get moving quickly and affect the areas that really matter to you, without requiring a huge time investment. Our Data Engineering Kickstart can help you get up and running quickly.

Book a free consultation →

Softwire