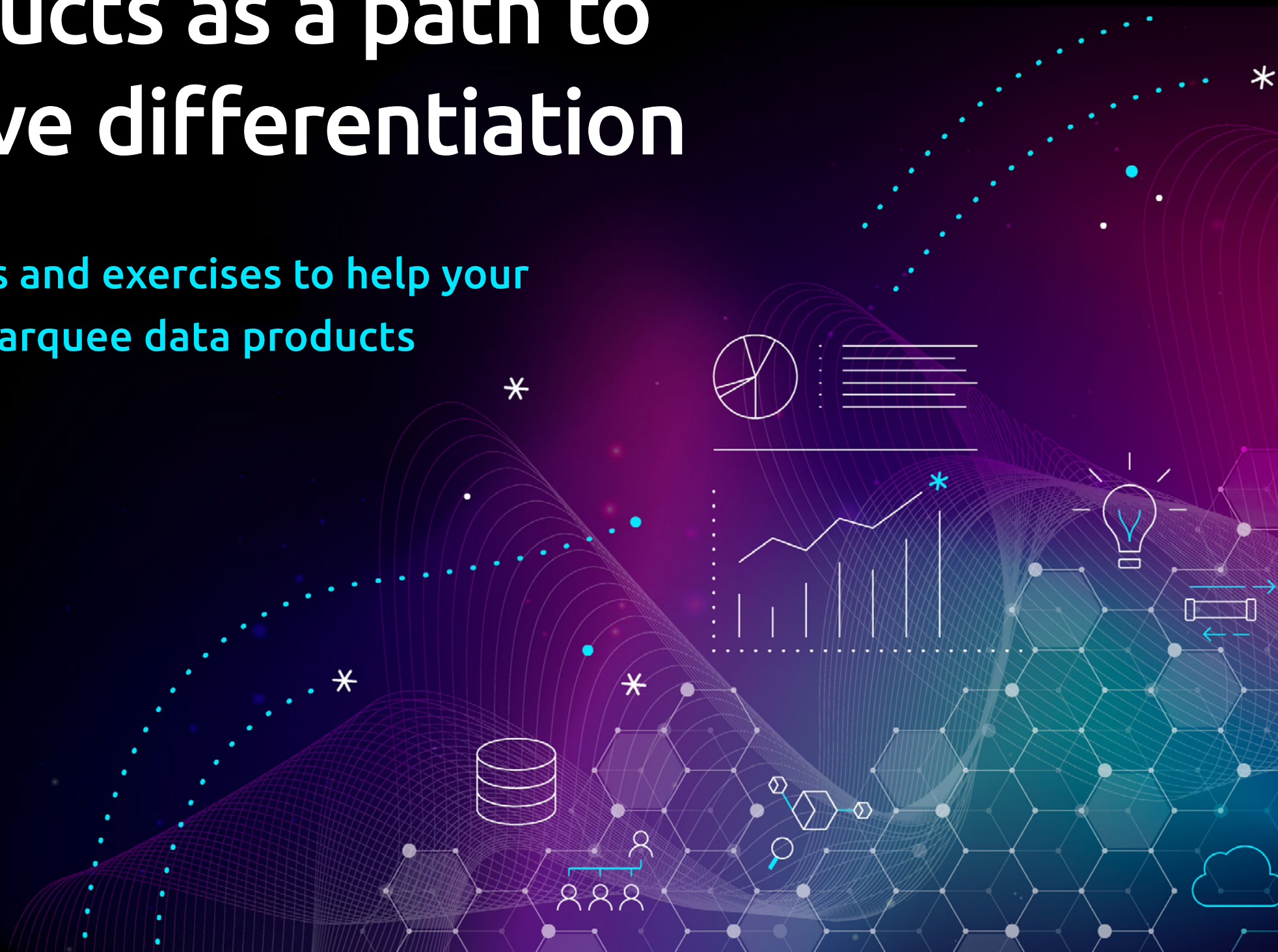




AN ACCENTURE EBOOK,
COMPLIMENTS OF STARBURST

Data products as a path to competitive differentiation

Six guiding principles and exercises to help your organization build marquee data products



Contents

01

What are data products 4

02

Data product guiding principles & exercise 8

03

Operations guiding principles & exercise 11

04

Evolution guiding principles & exercise 14

05

How companies are getting started with data products 17

06

Accenture x Starburst: data products as a bridge 19



Teresa Tung

Cloud First Chief
Technologist,
Accenture

**“Most people in the business
should care about data products
as a competitive differentiator.”**

Today many organizations are interested in data mesh as a means to create data products at scale and at speed to help the business make optimized decisions, innovate and better collaborate. Not only will data products assist with their digital transformation goals, but also AI transformation goals.


In a 2021 [analysis by Accenture](#), among executives of the world's 2,000 largest companies (by market capitalization), those who discussed AI on their earnings calls were 40% more likely to see their firms' share prices increase — up from 23% in 2018. We also project that AI transformation will take less time than digital transformation.



01

What are Data Products?

Data products are curated datasets packaged to create value for downstream consumers. This packaging makes it easy for consumers to use with documentation, service support, and may include applying business logic or analytics to data, combining data from multiple sources, and even generating synthetic data. To accelerate the creation and use of data products, we'll need the business to act as both a data product owner and the voice of the customer. In other words, the business really needs to be excited, care, and get involved.



As a way to demonstrate to the business what good looks like, start with a standout, marquee data product to demonstrate the value, set the standards and ways of working. Data teams are possibly one of the few within the organization that are excited about and understand the promise of the data mesh. As such, we need a simple case study that enthusiastic data teams can point to as an example as they evangelize data products, understand how to build and harness momentum, and get the business to care.

This initial data product then serves as a pattern that then can be fine-tuned by applying it to the many data projects that are already in flight across the business.

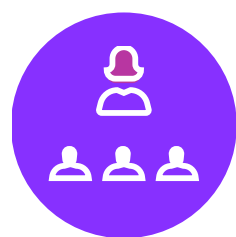
Identifying and deploying the right roles are critical to scaling data products. The organization needs to bring together technology teams and business domains from across the company. Traditional roles filled by the data teams should be increasingly filled by the people within the business who best understand the data and their own needs. This pivot starts with the data product owner, and eventually plays out in stewardship and implementation

roles as processes become better defined, self-service infrastructure becomes more available, and general data literacy improves across the business.

Today's data teams will move to become a design authority for data products responsible for the shared practices, processes, and tools that enable the business domains to create their own data products. The data team remains critical to incorporate the right level of participation across the business alongside defining and facilitating the organizational processes. For example, this team will help map those processes into the tools needed to build and manage data products. Even if the tools can vary and span across the business and even with partners, the central data team needs to certify their adherence to data standards to be safe, reliable, and efficient.

Use this initial data product to work out the roles, process, and initial tools to establish "what is a data product" and "how to measure its value," as well as assess organization and technology readiness. Then scan the list of ongoing and upcoming projects and systems to identify data mesh-ready projects that can be up-leveled into organized, managed, and valuable data products.

Below, we summarize the roles, processes and tools as foundational requirements for building your marquee data product.



Roles

Getting started requires getting the right roles as it will encompass how technology and business work together.

Some organizations start with the hub — a strong centralized team that handles a lot of the data. Focus on establishing a business-savvy and ideally business-sourced data product owner, possibly using a source aligned domain for simplicity.

Meanwhile, other groups might have very strong spokes with high data literacy across the company. And so these spokes might be already making different technology choices, different data product choices and have a lot of capabilities. Focus on establishing how domains work together and with the central team possibly using a cross-functional domain needing data from multiple sources to create novel value.



Processes

Organizational processes depend on the organization's structure. It is important to think through how heavy those processes need to be.

Organizations with lower data literacy might require a stricter, more prescriptive process in the beginning as business users begin to understand what it means to create data products and to work with data technologies.

For companies that have great data chops already, don't shy away from Agile ways of working under lighter-weight processes that enable new data teams to innovate frictionlessly.



Tools

Know that many tools are needed for data products and data mesh. Organizations may already have some in place, but there will be white space to fill.

One way to think about integrating tools is by looking at Agile software engineering, extending that idea to how data mesh applies to data.

With Agile, you have DevOps tools, processes and roles to guide a software product into production. In the same way, with data mesh, organizations need to figure out how to continuously integrate across the data supply chain and seamlessly deliver the data products.



Data Product


With the right roles, processes and tools in place, it will all be in service to create data products whose value can be measured. Keep in mind, cultivating a data product mindset is a challenge, yet it's vital because it brings value to a technology-led business. Technology alone can't create value by itself and so they need the business leader to be involved and define value – which can be measured in a multitude of ways including internal efficiency, new data-led solutions, direct monetization, and increased collaboration and innovation.

One thing we must get right is to sell this idea to a non-data business executive. When we talk about data mesh, it remains data geek speak. The question becomes, "How do I get a non-data executive to care?" A data product is something they could care about, but what does a data product even look like?

One proven method is to show them: take a part of the business that everybody cares about such as finance or HR, and create a data product. And these domains are oftentimes easier to start with because data is well understood and within their walls.

Most importantly, it's data that every part of the business needs.

This will be your marquee data product and your exemplar demonstrating what good looks like. In parallel, we'll work out how we support other data products and build out the foundation to scale. And when you hear something along the lines of, "I love that data product; I wish all my data was data products," you'll have the information that you need to drive your business decisions with competitive differentiation. Even better, you now have the playbook to move forward.



02

Data Product Guiding Principles & Exercise

In the following section, we'll take a closer look at a few guiding principles and, in parallel, reflect on a few questions as an exercise that you should look at with your organization. Naturally, the reflection exercise has a lot to do with the maturity of the organization and how much the organization can handle from day one and where you want to move towards.

Reflection Exercise: Is a data product defined narrowly or broadly?



Principle:

**All data products
should be designed
to be reusable assets,
vetted through
the authorization
framework**

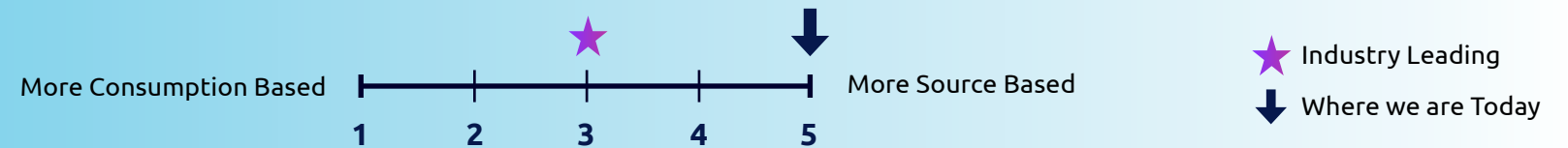
One of the most important things to start with is getting on the same page of what a data product is.

Many data products that are derived from existing data projects today might be easy to onboard but will be designed for pointed uses. Since data products should be designed to be reusable assets, these are some of the pivots that we might move away from data projects that are one-time use. However, let's not discount data products that are narrowly defined, so long as their value can be measured. Not all data products will be reused hundreds of times, but if there is utility in their data and they are packaged for reuse then they should be data products. The organization is already paying for those systems, so productizing their data is a no-regrets move.

If our goal may be to create more broader data products, then we can start with a marquee data product, one that makes sense with a broad user community that wants to support it. This establishes a pattern for moving towards more broadly defined data products that will have a wider definition of users.

No matter how you start, eventually, the vision is to have a mix. What's important is to establish a common understanding of what a data product is and how its value is measured.

Reflection Exercise: Are your data products more consumption-based or more source-based?



Principle:

The data strategy must evolve with the business strategy to drive competitive advantage

Your data products portfolio will likely contain a mix of both. Moving forward, we'll have consumption-based, we'll have source-based, and we'll also have aggregate-based.

In the beginning of your journey, it might be more source-based. The teams that are hosting some of these systems (and likely paying for these systems), might be a great place to start. Seek to understand why they're paying for the systems and the value they're getting from them.

Later data product entrants might leverage data from multiple sources — typically transactional data from one or more domains “aggregated” over a key master data dimension — in order to achieve their data mission. And after a substantial portfolio of foundational source-based and aggregate-based data products is built up, consumers may wish to “mash up” those elements to create novel value. These, if they are built with reusability in mind and according to the interoperability principles specified by the hub, are consumption-based data products, and they can be some of the highest-value pieces of the portfolio.

By answering the questions above, along with a data product mindset, it opens the door for more scalability and flexibility as businesses evolve with data-driven resiliency and competitive advantage.



03

Operations Guiding Principles & Exercise

In the next stage, if we know what the data product is, we might consider how we think about the creation of the product and the roles. Some organizations have more data capability and others might have less. Data might be still within the central data team who knows how to create data assets and manage data architectures. In this section, we'll take a closer look at a few Operations Guiding Principles and in parallel, reflect on a few questions as an exercise that you should consider with your organization.

Reflection Exercise: Who owns the data product?



Principle:

Data is a shared enterprise asset owned by the business and supported by IT

Today data is often a responsibility that's owned by IT. Meanwhile, with data products, we're striving for it to move it towards a business-owned asset. How far can an organization get? That depends largely on data literacy.

In the beginning, some organizations will still need the data team to play a lot of the roles such as the data product owner.

Initially, you might need to have that data team drive what it means to be a data product owner. "Mr. Business Person, could you help us really understand the use cases and the value from the data?"

The data team will need to help the organization understand the roles that need to be played and coach the business into fulfilling them.

Down the road, you should have the business owning the data product, owning the budget, and owning the product roadmap. Additionally, in order for the new owner to know the questions to ask — that will require training.

The good news is that data products are meant to provide the information needed for users across the business to get their jobs done. As data products usage expands, data literacy increases commensurately.

Reflection Exercise: Who leads data product definition and roadmap?



Principle:

The business will lead and champion the cultural shift required to scale

Today, data investments are largely IT-led. And as organizations move to investment and roadmapping that are business-led, we still require data and IT demonstrating how that's done, especially from a technology feasibility and effort perspective.

Organizations with a strong data hub will require clear definitions of its vision so that the business cares. And a value that shows the intended roles that the business should eventually take. The challenge will be adoption.

Meanwhile, organizations with strong spokes might have a lot of capabilities within their federated teams. But perhaps there's not enough incentive to work together. In this case, you might ask, "When there are needs for data connectivity to create a better aggregate product that is servicing many users, how do we actually bring these different domains or different parts of the business together?" The challenge will be interoperability.

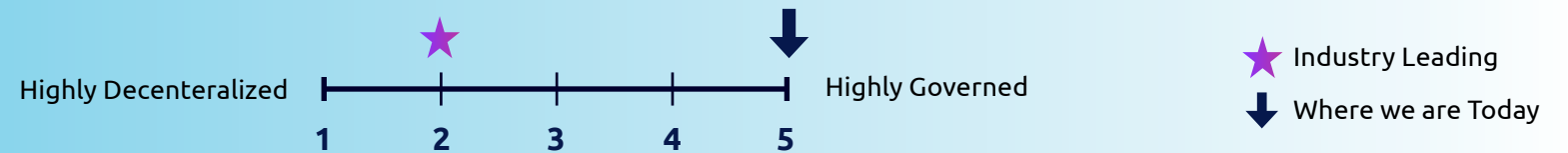
Part of operations guiding principles is how to make decisions on common data standards, process, and technology choices. This might mean standing up review boards or steering committees. These principles will likely still be guided by the central team to stand up that process, document the findings, and ensure that people are following what they agreed upon. But strong-spoked organizations should not reinvent the wheel and source proven, working principles from their domain teams.

04

Evolution Guiding Principles & Exercise

Finally, let's take a closer look at a few Evolution Guiding Principles as a way to involve and standardize data products as well as some exercises that will help you apply them.

Reflection Exercise: How much standardization is needed across your data products?



Principle:

Products with broader access require greater focus

Today, we might start with data products that are highly-governed, especially as we work through our initial products and processes. But the value comes from moving towards a system that is more decentralized.

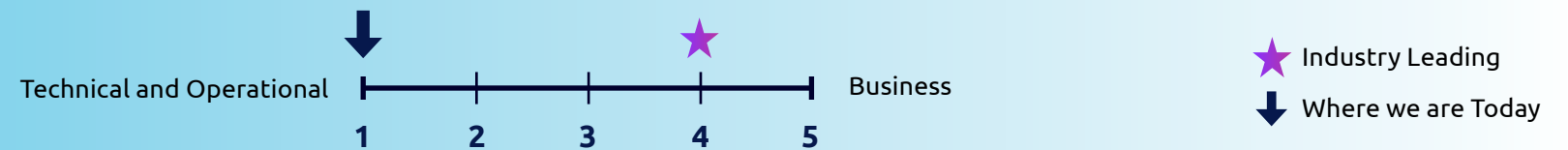
In the center, there are certain things we must have. For instance, we need security, privacy, and risk to be taken care of at the core.

But there are a lot of decisions about how much standardization is needed. How much adherence to industry ontologies? Is adherence across products within the company, organized by the domains, sufficient, or is cross-enterprise interoperability required? Which core products might be maintained by the central team for data protection? These choices become part of the business strategy, just like any other product.

In this way, the product-mindset becomes its own incentive: Better connected data products is a pathway to better products. Things that connect well with other products and within the industry — these will render a more competitive product.

As such, rather than requiring full standardization before it's useful to users, integrate incentivization as a means to create that standardization.

Reflection Exercise: What metadata is needed to evaluate and manage data products?



Principle:

Data is a shared enterprise asset owned by the business and supported by IT

Without metadata (that is, data about data), meaningful data products will be hard to find. And it's certainly impossible to manage data products to see what users are looking for, what products are used, which are used together.

Consider this analogy based on how internet search unfolded: Where we are with data products today is similar to when the world had AOL as its starting point to searching the internet.

At that time, AOL had curated services for sports, business, weather, etc. They built what they believed their users wanted.

Today, when we search on the internet for "Accenture", you get the location closest to you, our CEO, our stock price, and it's all packaged based on the context of the search. It's using metadata about your location, previous search queries and what you've found useful. And it's using metadata about the websites that exist — all to generate and package those search results cards we see.

And so, with the data products of today we might start with something similar to AOL. Let's curate and publish data products needed across the business.

This example shows what a data product is to its users. As more people use them, we can see how they evolve to create better products and more connected products. For that evolution we need the metadata so we could get to this place where it's like internet search today towards providing answers that better serve the end user.

05

How Companies are Getting Started with Data Products

As we're getting started with data products, consider these three different types of programs that you might already have at the moment. They offer various ways to get started on the journey.

Modernize/Migrate

Unlock Existing Data Products

Add a virtualization layer to unlock data products from existing data landscape. Often alongside data migration modernization program to identify and prioritize data products.

Accelerate

Pivot Data Org for Cloud Continuum

Transform centralized data organizations for federated governance over business-led domains. Define strategy and governance with associated federated tooling and processes. Industry-specific.

Grow & Innovate

Create Differentiated Data Product

Create highly differentiated data products unique in its interconnectivity, insights, privacy, quality (e.g. digital twin, customer 360). AI-powered data continuously evolves the product.

1

Unlocking existing data products.

Perhaps you have an ongoing cloud or data migration project. Or you might be modernizing systems. As you advance that work, find the data products already there. In addition, you might add a virtualization layer at the top of this migration or modernization in order to actually start seeing and shaping your data products. And then when the migration is done, you now also have this muscle that continues to allow you to create and experiment with new data products across systems and clouds.

2

Pivoting a data organization to cloud continuum.

So, when you're setting up governance strategies, that is a good place to add data product mentality to your end users. It's not just governance of data, it's governance of data products. It's not just a catalog of data, it's a catalog of data products. With this process, the end user and the role of the business domain are top of mind.

3

Creating differentiated data products.

Customer 360 and digital twin will work across many different domains and parts of the business. Figure out how to get these different parts of the business to come together in a way that gives folks control of the data that they have while creating this very differentiated data product that everybody has agreed is innovative and valuable.

“...when the migration is done, you now also have this muscle that continues to allow you to create and experiment with new data products across systems and clouds.”

06

 **accenture**

 **Starburst**

Data Products as a Bridge

Justin Borgman, Co-Founder and CEO of Starburst and Teresa Tung, Cloud First Chief Technologist at Accenture sat down with Dave Vellante of SiliconANGLE¹ and elaborated on how data products can serve as a bridge for your existing data investments. Teresa said, “I would love all my data to be centralized, mastered, and all within the same data warehouse... that’s just not the reality. The investment to actually migrate and keep that up-to-date, I would say it’s a losing battle. We’ve been trying to do it for a long time, nobody has the budgets, and then data changes, right? So you keep the data warehouse. I think it’s a very, very valuable, very high performance tool for what it’s there for.

But you could have this in a new mesh layer that still takes advantage of the things I mentioned, the data products and the systems that are meaningful today. And the data products that actually might span a number of systems, maybe either those that are source systems with the domains that know it best, or the consumer-based systems, or the products that need to be packaged in a way to be really meaningful for that end user. Each of those are useful for a different part of the business and making sure that the mesh actually allows you to use all of them.”

¹ SiliconAngle: [The Cube interview with Justin Borgman and Teresa Tung](#)

If you are moving towards adopting a data mesh architecture, we want to be there to help.

To learn more, visit the [Data Mesh Resource Center](#).

**Learn more about
Starburst Data Products**

starburst.io



How Starburst Data Products can help

Companies adopting a data mesh architecture must have an analytics engine capable of federating across these different data sources. Starburst is the analytics engine for the data mesh architecture, providing a single point of access to distributed data and empowering self-service analytics for each of the business domains.

With Starburst, there's no need to chase the idea of a single source of truth. Starburst is built on open-source Trino, a distributed engine that can execute SQL queries against data stored in a range of databases and file systems. With Starburst and Trino, teams can lower the total cost of their infrastructure and analytics investments, prevent vendor lock-in, and use the existing tools that work for their business so that they can concentrate on enabling faster time-to-insights. Trino's open technology means that integration with other open technologies such as data catalogs and data discovery tools is simpler and reduces the total cost of ownership of the self-service data platform.

Starburst Data Products is a module which allows data producers to create and maintain data products. This same platform can be used by data consumers to discover and understand and use these data products. Starburst makes it easy to produce, share, and consume data products using the built-in workflows and query editor. Starburst Data Products functionality includes a REST API, metrics, comments, incremental refreshes. The REST API allows engineers and developers to manage the complete life cycle of data products programmatically, with the scripting and automation tool of their choice. Creators and consumers of data products can add comments, and get a holistic view of metrics regarding the usage, profile, and consumption of all data products.

Starburst Data Products can benefit any analytics journey, large or small, on-premises or multiple clouds, in a warehouse or a lake. And federates across all of them.



About the authors



Teresa Tung

Cloud First Chief Technologist, Accenture

Teresa Tung is Accenture's Cloud First Chief Technologist and Chief Technology Officer for Data & AI and Edge. She leads the incubation and scaling of technologies in the cloud like edge, data mesh and heterogeneous infrastructure.

To learn more, visit [accenture.com/CloudDataAI](https://www.accenture.com/CloudDataAI) or [contact Teresa](#).



Scott Lee

Enterprise Data Architect, Accenture

Scott Lee is Accenture's Lead, Enterprise Data Architecture. He is an informatics specialist with a 25-year career in Data strategy, management, and engineering. Based out of Chicago, with experience in Americas and Europe on global-scale Data integration programs.

To learn more, visit [accenture.com/CloudDataAI](https://www.accenture.com/CloudDataAI) or [contact Scott](#).



About Accenture

Accenture is a leading global professional services company that helps the world's leading businesses, governments and other organizations build their digital core, optimize their operations, accelerate revenue growth and enhance citizen services—creating tangible value at speed and scale. We are a talent and innovation led company with 738,000 people serving clients in more than 120 countries. Technology is at the core of change today, and we are one of the world's leaders in helping drive that change, with strong ecosystem relationships. We combine our strength in technology with unmatched industry experience, functional expertise and global delivery capability. We are uniquely able to deliver tangible outcomes because of our broad range of services, solutions and assets across Strategy & Consulting, Technology, Operations, Industry X and Accenture Song. These capabilities, together with our culture of shared success and commitment to creating 360° value, enable us to help our clients succeed and build trusted, lasting relationships. We measure our success by the 360° value we create for our clients, each other, our shareholders, partners and communities.

Visit us at [accenture.com](https://www.accenture.com)

Building data products
as a competitive
differentiator

Accenture, 2022

