



The Top Data and Technology Trends for 2023 and Beyond

Industry experts weigh in on the trends that will define our industry

Table of Contents

Introduction	3
Data Democratization—Within Limits	4
Greater Focus on Data Governance	6
Increased Data Mesh Adoption	8
Data Contracts Enter the Scene	10
The Rise of Data as a Product (DaaP)	12
The Emergence of Data Architecture as a Service (DAaaS)	13
Data Modeling Makes a Comeback	15
More Data-Driven Automation	17
AI/ML: Beyond the Hype	19
Looking Ahead	21

Introduction

As 2023 is expected to bring a renewed focus on profitability over growth at all costs for businesses, we anticipate data will only become more important as a means of driving business value, customer loyalty, and continued innovation.

To break down where data is headed, we've compiled our own data predictions based

on analysis and insights from Coalesce.io executives, partners, and thought leaders across the data industry. From the return of data modeling and the growing importance of automation to more expansive data governance and democratization of data, on the pages that follow we've outlined the top trends in data for 2023 and beyond.

Contributors



ETHAN AARON
CEO & Co-Founder,
Portable



TONY DAHLAGER
Managing Director,
Analytics8



WAYNE ECKERSON
President, Eckerson
Group



MATT FLORIAN
Partner/Cloud
Analytics Practice
Director, COMERIT



KENT GRAZIANO
Owner and Chief
Strategic Advisor,
Data Warrior LLC



SATISH JAYANTHI
CTO and Co-Founder,
Coalesce.io



ARMON PETROSSIAN
CEO and Co-Founder,
Coalesce.io



CHAD SANDERSON
Chief Operator,
Data Quality Camp



MICHAEL TANTRUM
National Sales
Director, Resultant

Data Democratization—Within Limits

01

The push to make data available to a wider range of users will continue in 2023. “It has been something that people have been talking about and hoping would happen for at least the last decade,” said Michael Tantrum, National Sales Director at Resultant, a consulting firm focused on data analytics, technology, and digital transformation. “While it may have been called different things at different points, the goal has always been to get the data into the hands of key business users to make key decisions.”

The foundational idea behind data democratization is that data should be used by any number of employees and stakeholders to drive meaningful change within an organization, not just by those with a technical background. When team members—regardless of role—are enabled to work with data, the whole organization benefits. With information on their side, business strategy and decision-making can be rooted in reality, not based purely on experience or gut feelings.

“Helping teams who don’t have a data function get started with data is something I’m really excited to see,” said Ethan Aaron,

CEO and Co-Founder of Portable, an ELT platform that builds connectors on demand. Beyond creating opportunities for team members across an organization to use data, Aaron predicts data democratization will expand to include making the modern data stack accessible to more businesses in more sectors. “There are so many companies that haven’t transitioned from being an Excel spreadsheet or a Word doc company to being a data-driven organization,” he explained. “I believe there’s an opportunity to expand the addressable market by going out and showing people where they can eliminate manual processes with data, because every company has the opportunity to generate value from data. The problem today is that there is a high barrier to getting started.”

While there is undeniable momentum for data democratization, companies should support democratization initiatives mindfully. “It’s important that you give people the tools they need to answer questions for themselves,” said Tony Dahlager, Managing Director at premier data and analytics consultancy Analytics8. “But it is also important to have

organizational governance over the metrics
and the reporting that is cross-functional
in the business and widely used, in order to

maintain true consistency on definitions
and how things are being reported.”

**“It’s important that you
give people the tools
they need to answer
questions for themselves,
but it’s also important
to have organizational
governance over the
metrics and the reporting.”**

TONY DAHLAGER

MANAGING DIRECTOR, ANALYTICS8

Greater Focus on Data Governance

02

“Data governance means so many different things to so many people,” Tantrum said. For entities with large amounts of data such as health care, life sciences, and financial organizations, data governance expands beyond the walls of their business to include federal regulatory measures that address security and risk. If a business is small, data governance becomes more about documentation, shared definitions, and cataloging how things are built.

While definitions vary, data governance is, in effect, the management of data integrity in enterprise-scale systems, rooted in standards that are either internally defined or determined by government policies for data use and storage. As data democratization increasingly becomes the norm, data governance at every level is critically important.

“It’s no longer good enough to just build a data warehouse,” Tantrum said. “You have to add context and document how you built it. Data lineage, quality, glossaries—all these things are baked into good governance

because without them, how can people trust what you’ve built?”

Data governance took the spotlight in recent years amid the rise of new privacy rules and regulations, and is increasingly becoming part of every project’s DNA to create trust in the data and underlying systems. Trust is a major inhibitor to adoption, according to [IDC’s AI StrategiesView 2021 Survey](#).¹ As a result, sound governance is vital to fully unlocking data’s value.

“When I look at the economic market right now and into 2023, it’s not about accumulating new tools,” Aaron said. “It’s about justifying your data team’s existence from a cost and revenue perspective and how what you do leads to better strategic decision-making and true value.” Governance is a critical piece of creating those business benefits.

To meet the changing demands of regulations and maximize on data value, many companies are guiding their data governance programs by standing up new steering committees to work with engineers to define

standards for governance, implementation, and enforcement of processes that maintain legal compliance while driving growth and enhanced operations. We expect to see more conversations around governance in

2023. Their outcomes will likely affect data strategy at every stage of data collection, transformation, interpretation, and presentation.

“It’s no longer good enough to just build a data warehouse. You have to add context and document how you built it. Data lineage, quality, glossaries— all these things are baked into good governance because without them, how can people trust what you’ve built?”

MICHAEL TANTRUM

NATIONAL SALES DIRECTOR, RESULTANT

Increased Data Mesh Adoption

03

Already one of the hottest topics in the data world, we bet that data mesh will continue to dominate data conversations and see increased adoption throughout 2023. That's because as an organizing principle, data mesh gives businesses the agility and alignment they need to compete while also creating consistent standards even across decentralized teams.

Many companies are already integrating **data mesh concepts** into their businesses: having a centralized data team produce a foundational data layer that is then disseminated across departments and business users, who have a better understanding of the business rules for that data and are actually using the data. More will follow in 2023, even if some principles such as domain-first architectures, self-service, and data as a product (DaaP) are adopted a la carte, without larger data mesh programs being fully implemented.

"We are seeing organizations of all sizes considering adopting data mesh right now," Dahlager said. "It's no longer strictly on the periphery." As a sociotechnical paradigm and governance method, data mesh is so

relevant because businesses are increasingly focused on business value-driven data products as they mature and accelerate data investments. As a construct, data mesh pushes organizations to think of data in terms of use cases and quantifiable data—not just collecting data for data's sake.

"Data mesh aims to address the challenge of what to do when you've reached a plateau, when organizational complexity makes it impossible to get meaningful analytics out the door," Dahlager continued. For mature, data-driven organizations, data mesh is top of mind, although even smaller companies can benefit and grow using its core tenets.

In large and complex organizations, central teams often struggle to keep up with the demands of information. Data mesh is a new way to conceptualize how information is created from data, then managed and distributed to the organization to meet business demands. As such, we can expect data mesh to become more deeply ingrained in many business and application developments in 2023 and beyond.

Data mesh pushes organizations to think of data in terms of use cases and quantifiable data—not just collecting data for data's sake.

Data Contracts Enter the Scene

04

One simple, yet remarkably efficient way businesses can address data governance and data quality challenges is with **data contracts**, which are agreements between data producers and data consumers that consumers initiate to document data requests. A contract lays out the data itself, the schema, the SLAs, the business semantics, and other semantics of that data, and establishes ownership of that data. It's similar to an API, and there's a level of enforcement that ensures the constraints of the contract are actually followed.

Chad Sanderson, Chief Operator at Data Quality Camp, pioneered data contracts as a concept during his role as Head of Data at Convoy, and implemented them to solve the thorniest data quality issues the organization faced at the time. "Despite the fact that we were using the so-called 'Modern Data Stack,' pretty much everyone who was actually consuming the data was extremely unhappy with the state of our data warehouse," Sanderson said. "We had critical pipelines that were breaking on a regular basis, we had very little clear ownership. Massive tables had

evolved over years and years, and if those tables failed, there really wasn't a point of contact to own and fix the problem."

With data contracts, Sanderson in effect created a collaboration system where data producers and data consumers talked to each other about the data: what it would be used for and why it was valuable to the business. This not only increased transparency, but also created a better understanding among data producers on how the data pipelines they were responsible for would be used in the business. That, most notably, helped eliminate the need to involve data engineers whenever something broke and shifted ownership of the data back to the producer. "That's what I was trying to avoid, having the data engineering team be the bottleneck every time there was a problem with the data," Sanderson said.

In 2023, be on the lookout for data contracts to not only be a part of industry-wide conversations; they will likely enter more data strategy decks and implementations at organizations, especially as they look to adopt data mesh with a renewed focus on deriving business value from data.

“Data contracts are one way we can actually get a data mesh out and fully executed properly. It’s certainly something we’re very excited about here at Coalesce in helping companies evolve into. But there’s a lot of collaboration that takes place, there’s a lot of communication that needs to happen between the consumers and producers, and making sure that’s done in a way that’s properly enforced, properly governed.”

ARMON PETROSSIAN

CEO AND CO-FOUNDER, COALESCE

The Rise of Data as a Product (DaaP)

05

Forward-looking businesses are leaning into DaaP as a new revenue source. While organizations have long found value in their own data, DaaP takes advantage of the value data has to users beyond the business collecting it.

In a most literal sense, data as a product can be simply regarded as an opportunity to monetize data. “For the average business, it’s an extra revenue stream, because their data likely has value to someone else as well,” Tantrum said.

In 2023, DaaP will reach greater maturity, which will lead to more robust data organizations within enterprises. As companies increasingly leverage DaaP, the technological frameworks surrounding it may need to evolve. For example, companies may need to improve analytics functions or be more proactive in how they deal with the data they collect, through more refined governance and data modeling.

Suppliers, customers, and market research firms are all lining up to purchase data to enhance their products or services.

“It’s another one of those things that is driving direct revenue for businesses,” said Portable’s Aaron. DaaP is also an asset for data teams that must justify their relevance to big-picture value.

As companies seek to monetize data, they may need to evolve the technological frameworks that surround it. That may include more refined practices around data governance, data modeling, and more.

The Emergence of Data Architecture as a Service (DAaaS)

06

Coined by data industry veteran Wayne Eckerson, **data architecture as a service (DAaaS)** is a new paradigm emerging out of the decades-old practice of creating development templates to foster data reuse, accelerate data development, and improve data accuracy and efficiency.²

This has long been at the core of data architect roles. DAaaS commercializes this approach by using metadata to auto-generate code and documentation, as well as abstract the underlying platform so data architects and engineers can change or update the platform without impacting business users.

In 2023, DAaaS will enter the data landscape at scale, as it solves one of the biggest

problems the data industry has battled for years: data silos. DAaaS will enable more users across an organization to access and even build or modify data tables and pipelines, with built-in architectural guardrails that ensure compliance with data governance requirements.

“Dovetailing with data democratization, businesses are increasingly responsible for understanding and utilizing their sourced data to make informed business and technical decisions,” said Petrossian. “For that reason, we anticipate the rise of data-architecture-as-a-service tools to increase the quality, trust, documentation, and usage of data across enterprises. This is the only way a data mesh can be properly executed.”

“Dovetailing with data democratization, businesses are increasingly responsible for understanding and utilizing their sourced data to make informed business and technical decisions. For that reason, we anticipate the rise of data-architecture-as-a-service tools to increase the quality, trust, documentation, and usage of data across enterprises.”

ARMON PETROSSIAN

CEO AND CO-FOUNDER, COALESCE

Data Modeling Makes a Comeback

07

Data modeling has always been one of the most important steps in analytics projects, setting the groundwork to create databases, populate warehouses, manage data for analytical processing, and grant access to information in meaningful and governed ways. But it fell out of favor because it was “too hard and slow to do,” said Tantrum. Not surprisingly, the demand for speed often came at the expense of data quality. “The general sentiment was that we needed to get data to our users, so we pushed the product out, but all we did was push the problem downstream to users,” Tantrum added.

In fact, almost every expert we spoke with pointed to the return of data modeling as perhaps the most dominant trend of the year to come in data. “The thing is that data modeling has always been relevant, but it was just ignored,” said Kent Graziano, owner and Chief Strategic Advisor at consultancy Data Warrior LLC, and an advisor to Coalesce. Now, as businesses aim to extract measurable business value from their data, strategic attention has returned to data modeling.

Satish Jayanthi, CTO and Co-Founder of Coalesce, agreed: “Without a sustainable

data architecture to deliver high-quality data, businesses will not get very far toward the promise of predictive analytics, machine learning and AI, or even making truly data-driven decisions.”

More so, as businesses are preparing to face economic headwinds in 2023, the proper infrastructure can ensure they optimize compute costs, as well. “Data engineers are spending so much time discovering data that someone else had already scrubbed, cleaned, standardized—and not reusing data,” said Matt Florian, Partner and Cloud Analytics Practice Director at COMERIT. “And a data model really is about publishing reusable, consumable artifacts, which you can now reuse again. That will actually lower your compute costs because you’re not repeating pipelines.”

With a sound data modeling strategy in place, data can live up to its promises and deliver on the possibilities that come with sound practices, such as simplified, logical databases, reduced redundancies, minimal storage requirements, more efficient retrieval, and advanced analytics.

“A data model really is about publishing reusable, consumable artifacts, which you can now reuse again. That will actually lower your compute costs because you’re not repeating pipelines.”

MATT FLORIAN

PARTNER AND CLOUD ANALYTICS PRACTICE DIRECTOR, COMERIT

More Data-Driven Automation

08

Facilitating the rise of every other data trend in 2023 is the continued expansion of automation in functions across the data lifecycle. Given economic forecasts for the coming year, continued investment in automation can effectively help businesses do more with less, allowing for growth, productivity, and cost savings despite growing uncertainty.

Given that **Gartner has forecast hyper-automated businesses will reduce operational costs by 30% by 2024**, automation in data is a means of future-proofing.³ “In the modern data stack, we tell our customers to automate everything they can,” Dahlager said. “Good data engineers and contributors are hard to come by, so automation can help focus your teams on important and productive things that only they can do.”

While automation that tackles static processes may stall a bit in 2023, data-driven automation is different. ML applications can drive business transformation, creating resilience and measurable ROI, which makes them a key feature in tough economic periods.

“Anything pattern-based can be automated,” Tantrum said. “And so much of what we do in the data world can be automated.” He pointed to the “Six Ds” of automation that can transform common processes: design, development, documentation, deployment, daily operations, and data quality. “All these things can be done by machinery,” he explained. “We humans can augment them, but there’s no need to do it from scratch.”

Many experts describe automation as a predictor of competitive edge and innovation. As a result, 2023 will be a defining moment for enterprises that embrace digital transformation through more automated data processes. “In 2023, I expect to see a much greater adoption of automation tools to enable us to derive value from data faster,” said Graziano. “It will be a new era for data transformation and delivery platforms. The legacy ETL and ELT tools that got us this far will fall by the wayside as modern automation tools come to the fore with their simplicity and ease of use.”

30%

lower operational costs by 2024 for organizations that combine hyperautomation technologies with redesigned operational processes.

SOURCE: [GARTNER](#)

AI/ML: Beyond the Hype

09

Artificial Intelligence and machine learning have been at the forefront of data discussions for the past decade, but it's no secret that for many organizations, AI/ML initiatives remain an unfulfilled promise. Back in 2018, [Gartner predicted that through 2022](#) 85% of AI projects would fail, delivering erroneous outcomes due to bias in data, algorithms, or the teams responsible for managing them.⁴

We believe that was an accurate prediction, but given advances in automation across the modern data stack, we are optimistic that organizations will finally be able to advance their AI/ML projects to a maturity stage where they deliver tangible results for the business.

"As hyped as it was, machine learning and artificial intelligence is going to fundamentally change how the industry values data," said Sanderson. "In a world where some automated system is making a decision,

especially when that decision impacts someone in real life, be it in the medical or automotive space, the data is going to become critical."

Indeed, the biggest bottleneck in the AI and ML space in recent years has been the data governance that organizations need to implement to ensure that the decisions made—be it by a human or an algorithm—are accurate. "That's where companies are struggling right now and that's why we're so passionate about our mission to help make that data prep process as efficient as possible," said Petrossian. "Once organizations get to the point where they have their data properly prepared, persisted, with lineage and documentation in place, we can then get to this era of AI and ML and actually make decisions autonomously because you know that your data is accurate."

“As hyped as it was, machine learning and artificial intelligence is going to fundamentally change how the industry values data. In a world where some automated system is making a decision, especially when that decision impacts someone in real life, be it in the medical or automotive space, the data is going to become critical.”

CHAD SANDERSON

CHIEF OPERATOR, DATA QUALITY CAMP

Looking Ahead

We are in the renaissance period of the data-driven organization, which means the landscape is still being defined by pioneers in the space. Functionally, data and its associated technical concerns are now mainstream, while interest and enthusiasm for data's possibilities remain high. At Coalesce, we are honored to be shaping the future of the data landscape.

We believe 2023 will be a year of huge potential, when organizations of all sizes will continue to embrace new ideologies, processes, and tools to increase efficiency, reduce costs, and realize goals. These predictions for 2023 are reflections of today's data landscape and the values that guide our lives and work as engineers, strategists, and innovators in the data space. We can't wait to see what's next.

ABOUT COALESCE

Founded in 2020, Coalesce is the only data transformation platform built for scale. Coalesce combines the speed of an intuitive graphical user interface (GUI), the flexibility of code, and the efficiency of automation, empowering its customers with increased data team productivity and insights. Based in San Francisco, Calif., Coalesce is backed by Emergence Capital, 11.2 Capital, GreatPoint Ventures, and Industry Ventures and supports customers worldwide.

Learn more at coalesce.io.