



REPORT

The Top Data and AI Trends for 2024

Industry-defining ideas
for the year ahead



Table of Contents

Introduction	03
Contributors	03
1. Reckoning with AI's limitations	04
2. Ethical AI moves front and center	05
3. Data modeling is back, but now it's about semantics	06
4. Increased adoption of vector databases	07
5. AutoML becomes ubiquitous	08
6. Culture (and tools) evolve	09
Conclusion	09

Introduction

AI has touched nearly every industry and impacted culture on a wide scale over the past year. It has dramatically shifted how we think about, talk about, and plan for the future. In fact, AI touches nearly every one of our predictions for the year ahead.

Forrester estimates that investment in generative AI will grow from \$40 billion in 2022 to \$1.3 trillion by 2032, underscoring its momentum and dominance even at this nascent stage. As the industry adapts to AI, conversations about ethical frameworks, modeling, and how people participate in the process are coming to the forefront.

At Coalesce, we're eager to welcome a new year of innovation alongside our colleagues and partners, creating equitable and accessible data foundations. Welcome to 2024.

Contributors



ARMON PETROSSIAN
CEO AND CO-FOUNDER,
COALESCE



KENT GRAZIANO
OWNER AND CHIEF
STRATEGIC ADVISOR,
DATA WARRIOR LLC



MICHAEL TANTRUM
NATIONAL SALES DIRECTOR,
RESULTANT



SATISH JAYANTHI
CTO AND CO-FOUNDER,
COALESCE

Reckoning with AI's limitations

A pragmatic approach to gen AI

“The reality is that we’re still in the process of understanding AI,” explains Coalesce CTO and Co-Founder Satish Jayanthi. “We’re in the early days of this technology, and I think there will be some acknowledgment this year that we’re still working on understanding these models and their limitations.”

The buzz isn’t going anywhere, but in the new year, the realities of AI will eclipse the hype as businesses acknowledge the challenges. And there are quite a few, including costs. The AI programs that have dominated cultural conversations, such as OpenAI’s ChatGPT, Google Bard, and Synthesia, rely on a large amount of computing power to run queries and produce prompt responses. Organizations building models of their own are quickly learning that “getting answers with a certain latency isn’t easy—and it certainly isn’t cheap,” Jayanthi says.

But it isn’t just deploying AI that is costly—so is sustaining it. As demand for AI-driven applications increases, the advanced graphics processing units (GPUs) that power AI are in increasingly short supply. With constrained production capacity, AI development could become more costly or slow due to chip availability.

Considering the cost of development, hosting, and training models, we expect to see many new AI models shrinking in size in 2024 as developers trade scale for performance. They will focus on smaller enterprise models that emphasize sustainability and pay close attention to their supply chains. “These are the types of things we will see as we enter the second phase of AI, which will result in a more pragmatic approach to building,” Jayanthi predicts.



“The reality is that we’re still in the process of understanding AI. We’re in the early days of this technology, and I think there will be some acknowledgement this year that we’re still working on understanding these models and their limitations.”

SATISH JAYANTHI

CTO AND CO-FOUNDER, COALESCE

Ethical AI moves front and center

Focus on data lineage and governance

As we reckon with the limitations in building AI programs, we're also grappling with the underlying ethics and trust issues associated with how they are trained. "Data foundations are what make AI programs successful," says Coalesce CEO and Co-Founder Armon Petrossian. As such, lineage and governance are more relevant than ever.

"We have to be able to show where every answer we produce with AI comes from—not just as a one-off exercise, but as part of the DNA of our data programs," says Michael Tantrum, National Sales Director at Resultant. "Transparency and trust are becoming urgently important, because if you're training AI with bad data, it has the potential to be catastrophic."

Even the White House is weighing in on how AI should be developed and trained with recent [guidelines for the release, use, and deployment of AI to ensure protection for consumers](#). Last July, the Biden administration also called for third-party security checks and watermarking of AI-produced content to help users easily identify it. Yet, these remain voluntary protocols, and many worry that they don't also extend to the foundations of AI products.

Getting serious about guiding AI programs with quality data and consistent oversight isn't easy. Efforts must focus on implementation, anchored by concrete metrics. This requires creativity and collaboration on the backend, since there isn't a de facto method to measure something as nuanced as transparency. Yet, leaning into technical best practices offers a path forward. "We have to be deliberate about aligning people around it," suggests Tantrum. "As we find bad data in these systems, we then know who owns it and if it can be fixed. Secondly, automation ensures that every new piece of data acquired is tested and meets the expected standards."

Most organizations aren't asking *if* they should be employing AI, only *when* they will fully integrate it into operations. As such, they cannot afford to put conversations about ethics and data governance on the back burner. Translating questions of quality and lineage into concrete protocols and metrics allows companies to get ahead of scenarios that could be harmful to their customers, reputation, or bottom line.



"We have to be able to show where every answer we produce with AI comes from—not just as a one-off exercise, but as part of the DNA of our data programs. Transparency and trust are becoming urgently important, because if you're training AI with bad data, it has the potential to be catastrophic."

MICHAEL TANTRUM
NATIONAL SALES DIRECTOR, RESULTANT

Data modeling is back, but now it's about semantics

Creating a connection between natural language, the semantic model, and the technical database

Last year we predicted that data modeling would make a comeback, despite falling out of favor as enterprises prioritized speed, sometimes at the expense of quality. As AI continues to make its presence known in every industry, data modeling is roaring back into data programs. Defining how data is organized, stored, and the relationships within it, data models are the linchpin of high-performing AI. In 2024, these models will increasingly be semantic to meet the evolving needs of AI and large language models (LLMs).

“Semantic models use business and natural language to describe the relationships between things,” explains Kent Graziano, owner of Data Warrior LLC and a Coalesce advisor. “You need semantic models that have the right terminology so that regular people can use the resulting AI effectively. Then, you have to map that onto the physical database—that’s the technical part.”

As data volume grows, it is increasingly difficult for businesses to leverage it. Too often, only IT teams

are equipped to glean insights from data, creating more siloed environments with limited ability to turn data into actionable intelligence. Semantic models bridge the gap by mapping AI and LLM models directly to business value.

Knowledge graphs are one flavor of semantic modeling gaining popularity as they instantly render taxonomies and ontologies using data and relationships within a database. Knowledge maps are machine-readable data structures representing semantic knowledge relevant to a company, making that data more easily found and deployed by less technical team members. Knowledge graphs can improve the accuracy of LLM responses by as much as 52%, making it a worthwhile method to increase accessibility and performance.

“We have to create a connection between natural language, the semantic model, and the technical database to really take advantage of data,” Graziano says. “This is the learning curve we’ll see people grappling with in 2024.”



“We have to create a connection between natural language, the semantic model, and the technical database to really take advantage of data. This is the learning curve we’ll see people grappling with in 2024.”

KENT GRAZIANO
THE DATA WARRIOR LLC

Increased adoption of vector databases

The reason why searchable data is possible at scale

LLMs and AI are playing a key role in shaping the future of data management, spurring the rise of more readily searchable vector databases. “The demand for searchable databases has grown in tandem with a growing intolerance for things taking time,” Tantrum says. While the concept has been around and evolving for more than 30 years, searchable databases are increasingly possible due to the rapid adoption of AI and LLMs and their demand for vector databases.

Vector databases capture patterns in data and have custom algorithms for fast searching, producing results for exact or similar vectors. “Vector databases are another evolution of database technology. We have relational databases, we have NoSQL databases, we have graph databases, and now we have vector databases. They are purpose-built for AI and LLM applications and are the sole reason why searchable data is possible at scale,” offers Jayanthi.

Traditional databases have limitations for handling unstructured data—which is most data generated

today. Vector databases, however, are equipped to handle large and complex data types effectively, with advanced indexing and search algorithms for more efficient similarity searches, helping to establish relationships between words, phrases, and sentences with understanding of how they are related and different. Those relationships allow models to match existing data to any query, conjuring directionally correct answers. They make LLMs practical and possible by allowing users to avoid running every new query through a training model.

Databases tend to evolve to meet the needs of the current technology. Vector databases have evolved to meet the needs of AI, marrying unstructured data sets and the power of LLMs. 2024 will demand greater understanding, access, and deployment of vector databases to match the expansion of AI systems. As a driver of AI’s accuracy and usability, these databases will become omnipresent for data management in the new year.

“2024 will demand greater understanding, access, and deployment of vector databases to match the expansion of AI systems.”

AutoML takes over

Tools like Copilot become ubiquitous

Another 2023 prediction we got right last year:

Automation is indeed becoming ubiquitous in our industry.

Automated machine learning (AutoML) is trending upwards as the demand to make AI accessible grows. AutoML allows less-technical users to train, optimize, and deploy models quickly. With an anticipated CAGR of 43.9% between 2023 and 2028, AutoML is expected to grow into a nearly \$8 billion market within the next five years.

Taking over the routine tasks of data scientists and developers such as data preprocessing, model selection, tuning, and optimization, these tools are advancing rapidly to create models with better efficiency and scalability. They may even replace how AI and ML applications have conventionally been developed up to this point.

“I think tools like Snowflake Copilot are going to become ubiquitous in the next year,” says Petrossian. The solution can generate SQL based on an input question via LLM technology. “Tools like this are making it possible for businesses to automate previously very technical tasks and get more value out of their data,” Graziano adds.

As AutoML creates greater efficiency, tools will proliferate. Businesses need to assess how their teams will use AutoML tools and their data science expertise before rushing to adopt them. Companies also need to determine how these programs can most effectively support existing tech stacks and business intelligence needs.



Expected size of the AutoML market in the next five years

Source: Mordor Intelligence

Culture (and tools) evolve

People are more important than tools and technology

6

For a long time, the IT industry has focused on developing tools that can unlock the possibilities of automation, AI, and ML. Now, we have entered a new era of IT tooling and are rethinking people's roles within data programs. "People are now more important than tools," Jayanthi says. "How we structure data programs and their people is a new paradigm shift—one that might take some time to become evident."

Data tools now allow less technical business users to get closer to their enterprises' data foundations. Simultaneously, as the benefits of data programs expand to touch every part of business, new demands are being put on IT teams. "The world is more collaborative," Tantrum says. "As tools take care of the engineering aspects, we need people who can think creatively and communicate effectively. It's a different set of skills than we've required data people to have before."

"This has a lot to do with changes in the whole ecosystem," Graziano adds. "Business analysts are working with databases and the whole culture is shifting to doing things in business semantic terms. That's why it is so important to model the data in business terms and model the concepts of the business. You can have all the tools in the world, but if you don't have people who can translate between business and IT, it won't work."

As we enter 2024, we will see all sides of the enterprise upskill. The business side will continue learning the tools that bring them closer to the data, while IT teams expand their ability to communicate and collaborate with business stakeholders. This shift toward deeper collaboration is the key to meeting today's demands for accelerated insights, efficiency, and problem-solving that lead to innovation.



"People are now more important than tools. How we structure data programs and their people is a new paradigm shift—one that might take some time to become evident."

SATISH JAYANTHI

CTO AND CO-FOUNDER, COALESCE

Conclusion

AI is an increasingly important part of both our tech stacks and everyday lives. The trends and conversations discussed in this report aren't just happening quietly in IT departments, they are an unprecedented part of the national discourse. It's an exciting time to be working in a space that is rapidly evolving and influencing a technology that is changing the world.

In 2024, we must ensure the data underpinning AI programs is trustworthy and equitable. We must also lower the barriers to entry for this transformative technology and improve our ability to collaborate and solve problems. All of us at Coalesce couldn't be more excited to continue this work alongside our brilliant partners—and we can't wait to see what's next.



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 supports customers worldwide and is backed by Emergence Capital, 11.2 Capital, GreatPoint Ventures, and Industry Ventures.

[Learn More](#)



© 2023 Coalesce Automation. All rights reserved. [Coalesce.io](https://coalesce.io) | sales@coalesce.io