snowflake®

DATA TRENDS 2024
FINANCIAL SERVICES

How Industry Leaders Are Building for Success in the Snowflake Data Cloud

TABLE OF CONTENTS

GEN AI IS TRANSFORMATIONAL

In 2023, the hype surrounding generative AI was massive. But in 2024, the real work is underway as enterprises start to pour their proprietary data into large language models (LLMs) and create bespoke applications that drive their businesses in new directions.

At Snowflake, we see this trend clearly, including among some of the largest enterprises in the Global 2000. Gen AI has moved out of the experimentation phase and into production. Our recent report, **Snowflake Data Trends 2024,** reveals how organizations have been using this transformative technology and other data-driven workloads within the Data Cloud from January 2023 to January 2024. Here are the topline results across industries:

- **The number of active accounts adopting machine learning (ML) functionality within Snowflake increased 67% since the ML functions of Snowflake Cortex went into public preview in June 2023. That opened up more possibilities because data scientists and other experts are no longer a bottleneck.**

- **Organizations' usage of data governance measures is rising, and seems to be improving their ability to use data: While usage of tagging data rose 70% to nearly 100%, the number of queries against protected objects is up 142%.**

- **The number of native apps created by Snowflake users increased by 311%, while adoption of these apps soared 96%.**

- **Usage of Python, a very popular language for AI development, grew 571% — considerably more than any other language year over year.**

In this report, we'll take a closer look at how Snowflake user accounts are working with their data in the financial services industry — one of the earliest industries to adopt AI and ML — to see how generative AI and other data-driven technologies are being deployed across this sector, and explore the factors influencing AI adoption in financial institutions.

Our research reveals how financial services organizations are interacting with the tools and features within the Snowflake Data Cloud. Understanding these trends and what they signify can provide valuable clues to inform industry leaders' business strategies, planning and technology investments over the coming year.

METHODOLOGY

We looked at how our Snowflake financial services customers adopted features and capabilities of the Data Cloud over the previous fiscal year to reveal trends, both in terms of the foundational development of data infrastructure and those users' first moves into advanced AI. Where relevant, we have compared industry usage to our broader, cross-industry metrics to show both alignment and, more importantly, deviations. Generally, we compared usage in January 2023 to January 2024 to align with Snowflake's fiscal year, except in cases where features went into public preview during the year. In those cases, we compared the first full month in public preview to January 2024. For the full methodology, **see the appendix.**

GEN AI FOR FINANCIAL SERVICES:
THE REGULATORS ARE COMING

Financial services is among the most heavily regulated industries. Even so, compliance issues have become especially acute with the emergence of transformative technologies such as generative AI, for which the rules are still being written. Over the past year, several major initiatives have been put into place to establish guardrails for the use of gen AI in the financial sector.

In July 2023, for example, the Securities and Exchange Commission (SEC) proposed **new rules for broker-dealers** to mitigate potential conflicts of interest stemming from the use of predictive analytics. Among the SEC's concerns are potential conflicts of interest or bias within AI-generated trade ideation or investment planning that do not meet the requirements for best execution.

In December 2023, lawmakers in the European Union agreed on a set of principles **governing the use of AI** across a wide range of use cases. Requirements relevant to financial services firms include a call for greater transparency into the data used to train AI models, as well as the implementation of **risk assessments for high-impact applications** of AI.

The **White House Executive Order on Artificial Intelligence,** released on Oct. 30, 2023, gave the U.S. Treasury Department 150 days to issue a report laying out the best practices for financial institutions to manage AI-specific cybersecurity risks. New rules could appear as early as the end of March 2024. In Congress recently, U.S. legislators introduced **no fewer than 30 bills** addressing bias, reliability, transparency, security and other issues surrounding generative AI, with **more likely to come.**

The heightened attention regulators are paying to AI makes effective data governance more crucial than ever. And that is consistent with our findings in our **Data Trends 2024 report.**

Financial institutions have long required a strong data foundation to operate successfully. Now, in an era defined by AI, this tenet may be the single most important factor for success. Paramount to that success is a unified data platform that eliminates internal silos while enabling seamless access to structured, unstructured and third-party data. To reach this state of operational effectiveness, robust governance mechanisms must be in place to minimize data leaks or unauthorized access, as well as to adhere to regulatory demands.

FIVE CONCERNS SURROUNDING GEN AI

Generative AI's creative capabilities are groundbreaking. This transformative technology is likely to have a profound impact across all sectors of the economy, especially for financial services. But gen AI advances also come with **a number of potentially serious risks.** Here are the the most significant areas of concern:

1 **Unintentional bias:** All ML models contain the potential for bias in their training data. For example, under- or over-sampling of a particular population has been shown to skew the results of facial recognition algorithms, adversely impacting some demographic groups.

2 **Lack of transparency:** In many cases, the training of data for a particular model is unknown or undisclosed, making it difficult to determine if the data is biased or based on unreliable sources. In addition, the neural networks used to build LLMs are typically "black box" applications whose inner workings are opaque even to the people who designed them. The ability to demonstrate how an AI model makes decisions — especially those that have a potentially adverse outcome for customers — is likely to be addressed by future regulatory activity.

3 **Data leaks:** Public-facing gen AI chatbots have leaked source code and other personal or proprietary information that was uploaded by users. These bots may also inadvertently reproduce materials used in their training data — a flaw known as "regurgitation."

4 **Inaccurate or false results:** AI chatbots sometimes generate fictional information in response to queries — commonly called "hallucinations." Although new iterations of popular bots have significantly reduced the number of hallucinations, their results are not ready to be taken at face value.

5 **Cybersecurity issues:** As with all technologies, bad actors are using AI to deploy new exploits and attack AI chatbots themselves. The ability to corrupt the data used to train an AI application (known as "data poisoning") or use carefully constructed prompts to extract sensitive information (referred to as "prompt injection attacks") are two techniques bad actors now employ.

TRENDS THAT MATTER
TO FINANCIAL SERVICES

Financial services organizations aren't just embracing data analytics and AI tools with greater frequency, they're also using data in different ways and at different volumes than we're seeing in other industries. Our research identified three significant trends that show how data modernization is having a foundational impact on financial services organizations, influencing everything from business strategy to planning operations and technology investments.

TREND ONE:

ENTERPRISES ARE TAPPING INTO UNSTRUCTURED DATA

Unstructured and semi-structured data make up **an estimated 80% to 90%** of all enterprise data, and the volume continues to grow. The financial services industry houses a high volume of unstructured data, encompassing PDFs of loan agreements and insurance policy applications, emails with clients or business partners, call center recordings and more.

The processing of unstructured data increased 123% across the Snowflake Data Cloud since the features in Snowpark became publicly available on June 27, 2023. But within the financial services industry, that figure was 244% — nearly twice the overall rate.

One reason for this dramatic increase may be that the financial services industry inherently produces more unstructured data than other industries, because of record-keeping requirements particular to the industry. Beyond that, though, is the fact that financial services firms are eager to unlock the value in such data to identify new customer segments, drive better services and streamline call center operations. Tapping into the power of unstructured data will not only accelerate advances in 360-degree customer views, marketing analytics, fraud detection and investment research use cases, but will also enable generative AI applications including the creation of gen AI copilots.

The processing of unstructured data in financial services increased 244%, nearly 2x the overall rate
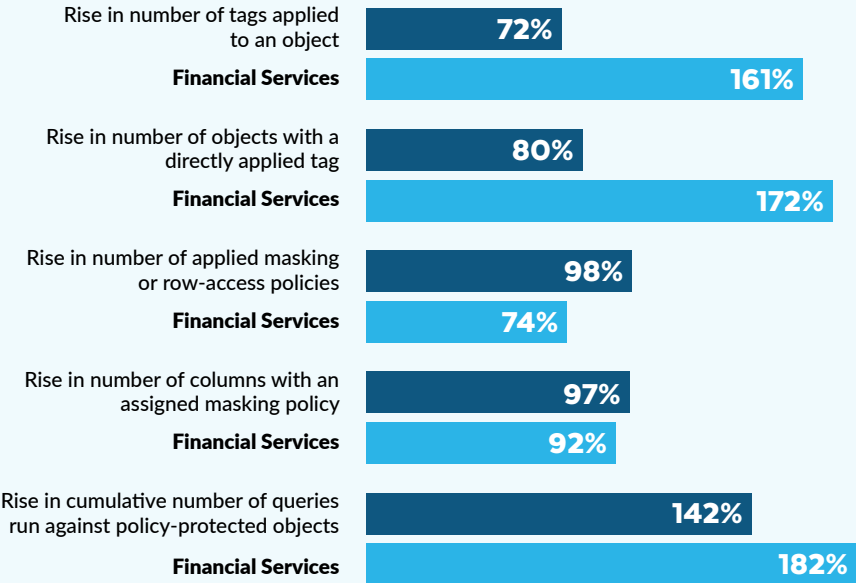
DATA GOVERNANCE IS GETTING MORE REFINED

Efforts at data governance are rising across the board. In the Snowflake Data Cloud, this is evident in the increased use of tagging (for example, to identify sensitive policy-protected data) and masking (to restrict access to policy-protected data). In general, the use of data governance measures rose 70%–100% in the last year.

Still, these added governance measures are not preventing organizations from unlocking the value contained in this data. Overall, the number of queries for access to policy-protected data assets rose 142%

across all industries over the last year, indicating that organizations are making deliberate, approved and proper use of their data. And over the same period for financial services, the increase was even greater, coming in at 182%. This can be attributed to the fact that financial services companies are both exceptionally data-driven and have the experience, motivation and skills required to do it more effectively than organizations in most other industries. This should prove extremely helpful as financial firms evolve their use of AI.

As the graph below shows, the financial sector relies more heavily on tagging than masking, and came in a little below the general growth rate on applied masking and row-access policies, as well as masked columns. But the industry increase for tags applied to data objects was significantly over the average (nearly double the growth rate). This is most likely a result of previous governance efforts and well-established access policies, spurred by government compliance requirements. Historically, financial service firms have used tags to protect data more than other industries.

ADOPTION OF GOVERNANCE FEATURES IS ON THE RISE

Rise in number of tags applied to an object	72%
Financial Services	161%
Rise in number of objects with a directly applied tag	80%
Financial Services	172%
Rise in number of applied masking or row-access policies	98%
Financial Services	74%
Rise in number of columns with an assigned masking policy	97%
Financial Services	92%
Rise in cumulative number of queries run against policy-protected objects	142%
Financial Services	182%

TREND THREE:

PYTHON USAGE IS UP DRAMATICALLY ACROSS THE BOARD

The Python programming language is considered particularly well suited to AI programming and building AI apps. So it's no surprise that it was by far the most commonly used scripting language in Snowpark (our coding library component of the Snowflake Data Cloud).

Across all industries, Python use grew 571% in the last fiscal year, easily outpacing Scala (up 387%) and Java (up 131%). Comparatively, for financial service customers the use of Python shot up 585%.

Since organizations are increasingly strategizing around and experimenting with LLMs and generative AI, it's no surprise we're seeing considerably more work being done in Python. Beyond that, for developers in the financial services industry Python has been the dominant language for years. Python's open source roots and greater accessibility make it a better choice for some organizations than more statistically based programming languages, such as R, MATLAB or SAS.

These findings suggest that IT teams in the financial industry are in a good place to launch advanced AI initiatives, since they've already oriented their teams around Python. Talent is always in short supply in IT, but because so many developers working in the finance industry are already deeply familiar with Python, they should be well positioned for gen AI development.

PYTHON USAGE IN FINANCIAL SERVICES INCREASED

585%

THE STATE OF THE FINANCIAL SERVICES DATA CLOUD

Financial services firms account for a significant portion of Snowflake's Data Cloud, which is uniquely suited to the industry's needs.

Indeed, for years many of the world's leading financial institutions have turned to Snowflake for help creating 360-degree views of their customers, streamlining risk management processes, improving their threat detection, minimizing reputational and financial risk from fraud, and adhering to regulatory guidelines.

The need for a unified, secure, cloud-based platform that can handle vast amounts of data continually increases as developers put more AI applications into production and bring more AI/ML work into Snowflake. Fortunately, Snowflake was built to address this need. It allows firms to scale multiple workloads across their entire data stack using a single copy of their data, enabling a single source of truth without requiring the organizations to extract, transfer and load information from an existing data warehouse or data lake. And **Snowflake Marketplace** provides access to over 600 providers and 2,416 live, ready-to-use data sets, services and Snowflake Native Apps (as of January 31, 2024).

Meanwhile, regulatory demands and business resilience needs require financial firms to spread their data across multiple clouds. Since Snowflake works seamlessly across multiple providers, organizations can easily use separate clouds to collaborate for different purposes. This gives companies faster access to the global financial services ecosystem of payment processors and application providers. A multicloud approach is also advantageous when it comes to building business resiliency and better disaster recovery. And because financial services companies can share data entirely within Snowflake, they can exchange data more quickly and cost effectively.

Since security and governance are key priorities for financial services firms, strict controls are required in these areas. Snowflake eliminates data replication by allowing data sharing entirely within the platform, so it's more secure than legacy data-sharing methods such as FTP, SFTP or email.

In a rapidly transforming business environment, the **Snowflake Financial Services Data Cloud** offers several key advantages:

- **Customizability:** Financial services organizations can customize generative AI and LLM solutions within hours, and easily and securely use gen AI and LLMs in seconds, including third-party commercial LLMs when they're determined to be the best approach. These solutions can also be kept in-house, further enhancing organizations' ability to govern policy-protected data and safeguard their intellectual property.

- **Accessibility:** With Snowflake, companies can grant access to insights derived from data to all authorized users within the organization, democratizing data-driven decision-making beyond just data scientists and other AI experts. People in any part of the business can write code thanks to gen AI integration, and answers to their critical questions may be just a click or query away.

- **Scalability:** The unified, easy-to-use Snowflake platform allows financial service organizations to build a connected data ecosystem of leading data and app providers, top solution partners and critical service providers. This enables them to power many varying workloads at scale.

By simplifying access to data, the Snowflake Financial Services Data Cloud makes it easier for banks, asset managers, insurers, payment processors and other intermediaries to collaborate securely. That leads to faster decision-making and improved competitiveness.

With operational and IT burdens alleviated, enterprises can do more with less and focus more on development. Snowflake also equips companies to deliver custom applications faster and more securely than ever by keeping their most precious assets — their data and the models they create from it — close at hand.
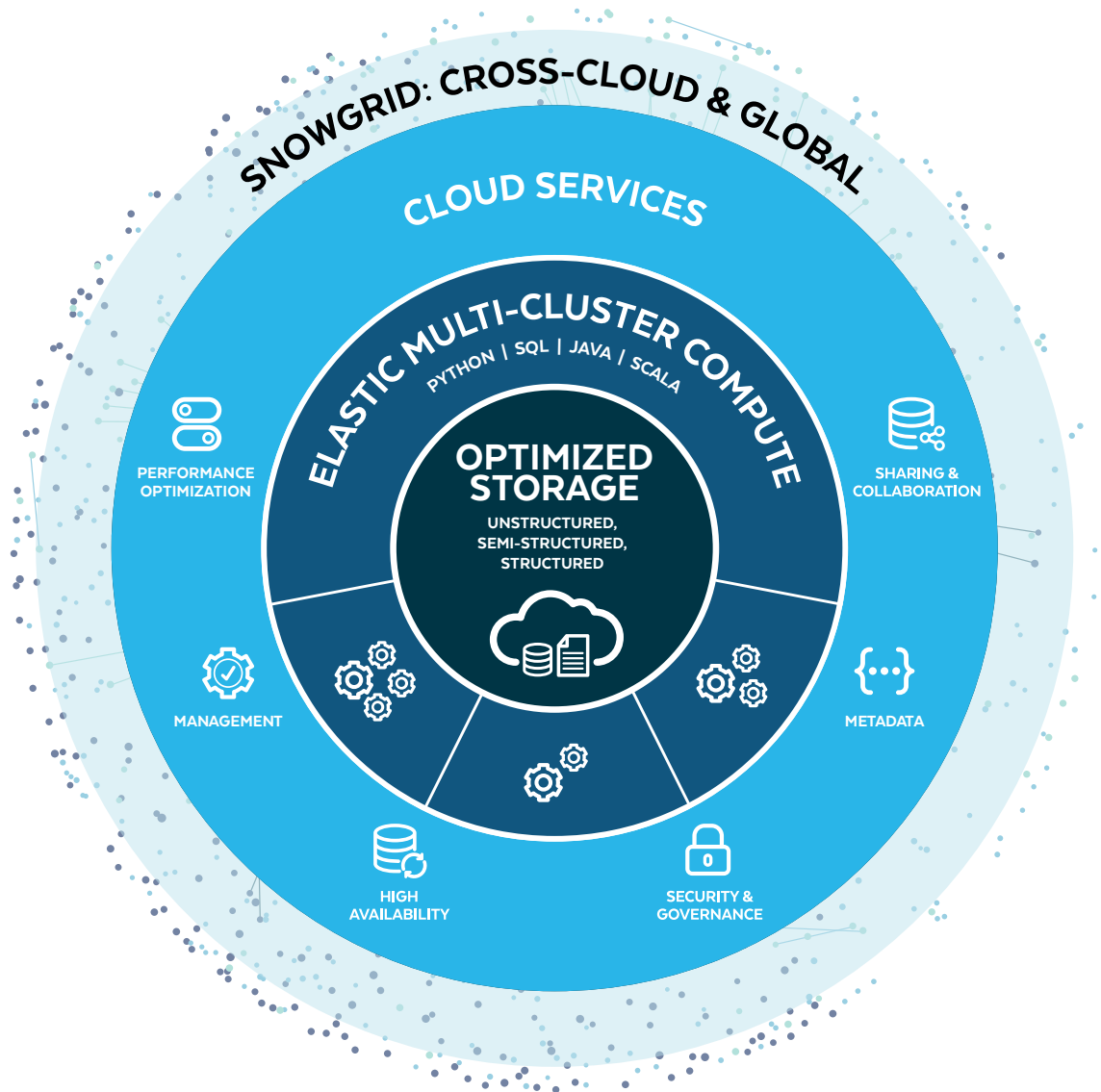
Snowflake Platform Architecture

USING DATA SHARING AND COLLABORATION
TO SOLVE CRITICAL FINANCIAL SERVICES BUSINESS CHALLENGES

The Snowflake Data Cloud modernizes how organizations access and leverage data via secure data sharing and data collaboration. And with **Snowflake Marketplace,** financial services companies can access live, ready-to-use data, services and Snowflake Native Apps. Here are three key areas where organizations are using data sharing and collaboration in financial services to gain a competitive edge and solve critical business problems:

Investing, portfolio management and quant research

For financial services firms to speed research cycles, turn quant-driven investment analytics into insights and ultimately get ahead in the market, they need access to ever-increasing amounts of further differentiated data. That's why data acquisition and ingestion are so critical for investment, portfolio management and quant research data workflows.

The ready availability of such data sets helps quant researchers, portfolio managers and traders alike uncover market-differentiating investment and trade opportunities. Here are some Snowflake Marketplace providers that offer these capabilities:

- **S&P Global:** S&P's complete view of global energy and commodities markets allows customers to make decisions with conviction and create long-term, sustainable value. Vital to navigating energy transition, S&P Global Commodity Insights' coverage includes oil and gas, power, chemicals, metals, agriculture and shipping.

- **FactSet:** An industry leader in acquiring, integrating and managing financial data, FactSet's content gives organizations the power to monitor global markets, research public and private companies, and gain industry-level insight to ensure businesswide consistency, transparency and data integrity.

- **LSEG:** As one of the world's leading providers of financial markets infrastructure, the London Stock Exchange Group (LSEG) delivers financial data, analytics, news and index products. It helps organizations fund innovation, manage risk and create jobs by partnering with customers at every point in the trade lifecycle.

- **ICE:** Intercontinental Exchange (NYSE: ICE) is a leading global provider of data, technology and market infrastructure. The NYSE is the premier global venue for capital raising. NYSE Arca Options and NYSE Amex Options are leading equity options exchanges.

- **Visible Alpha:** Backed by the world's leading investment banks, Visible Alpha is dedicated to creating the deepest consensus in the market, with more and higher-quality sources, and longer-term forecast horizons. It brings clarity to the market's complex spectrum of ideas, perspectives and analysis.

Sales and marketing

Driven by the push to better target, retain and grow customers, financial services firms are using data sharing and collaboration capabilities in the Snowflake Data Cloud to maximize customer data acquisition, analytics data capture, measurement and attribution. Below are some key Snowflake partners that financial services companies are leveraging for sales and marketing use cases:

Analytics data capture

- **Data.ai:** As the market leader and trusted data source for the app and digital market, data.ai empowers customers in growing their apps, investment portfolios and digital business with its proprietary data and insights. Its customers include some of the world's leading technology companies across finance, media and gaming.

- **Funnel:** Providing business-ready data that is clean, accurate and always up to date, Funnel helps marketers become more data-driven. It allows organizations to collect data from all platforms, and then transform data with recommended and customizable rules, and explore it to find new insights.

- **Heap:** With Heap, organizations of all sizes can remove technical bottlenecks and gain a single, comprehensive view of their customers through software that automatically collects, organizes, analyzes and connects customer data. This lets businesses create more valuable products and experiences, discovering insights and taking action instead of building pipelines or tagging manually.

Customer data acquisition

- **HubSpot:** HubSpot is a leading customer relationship management platform that provides software and support to help companies grow better. The platform includes marketing, sales, service, operations and website management products that start free and scale to meet customers' needs at any stage of growth.

- **Braze:** The Braze customer engagement platform delivers messaging experiences across push, email, apps and more. With data, technology and teams working together in unison, the platform allows for real-time and continuous data streaming, replacing decades-old databases that aren't built for today's on-demand, always-connected customer.

Measurement and attribution

- **NielsenIQ:** The world's leading consumer intelligence company, NielsonIQ delivers the most complete understanding of consumer buying behavior and reveals new pathways to growth. With a holistic retail read and the most comprehensive consumer insights, it delivers advanced analytics through state-of-the-art platforms.

Risk management

Across the financial services sector, risk and compliance teams continue to grow in size and scope, as many look to minimize regulatory, reputational and revenue risks that might follow noncompliance. Areas of risk management include financial and transactional threats such as liquidity and credit risks, but also those related to ESG and sustainability, as well as fraud and other financial crimes. Here are several partners that Snowflake data users are leveraging for risk management:

- **Cybersyn:** Cybersyn, a Snowflake affiliate, is a DaaS (data-as-a-service) company, whose mission is to make the world's economic data transparent to governments, businesses and entrepreneurs, and enable a new generation of decision-makers.

- **MSCI:** A leading provider of critical decision support tools and services for the global investment community, MSCI powers better investment decisions by enabling clients to understand and analyze key drivers of risk and return and confidently build more effective portfolios. Clients use its research-enhanced solutions to gain insight into and improve transparency across the investment process.

- **Black Knight:** Black Knight's industry-leading data and analytics solutions help clients reduce risk, satisfy regulatory requirements, improve retention and create opportunities for growth. Mortgage, real estate, capital markets and other professionals across the U.S. leverage Black Knight's comprehensive data and innovative analytics to help inform their critical business decisions.

- **CoreLogic:** CoreLogic provides information intelligence to identify and manage growth opportunities, improve business performance and manage risk. A market leader for unique, property-level insights backed by science and analytics, CoreLogic delivers value across its clients' business operations, addressing challenges and acting quickly to present innovative, cost-effective solutions.

FIVE WAYS FINANCIAL SERVICES FIRMS
CAN LEVERAGE GENERATIVE AI

Though it's still ramping up, generative AI already promises a wide range of benefits and potential use cases for financial services firms. Here are five ways that these organizations can benefit from this transformative technology:

1. Increased productivity and an improved customer experience in banking

According to recent research, the banking sector has a high potential for gen AI-powered automation when compared to other industries. The research found that 73% of the work at banks could be fully automated or augmented to make employees' work more efficient. For banks with call centers or those looking to enhance customer service, gen AI can help by summarizing call transcripts, creating new customer profiles, assessing customer sentiment, and implementing other automations to reduce costs, improve customer experiences and ultimately retain valuable customers.

2. Faster investment analytics and quant research in asset management

For portfolio managers and quant researchers, the ability to quickly scan through unstructured data within thousands of analyst reports, news and other documents, while also bringing in third-party market, reference and company data, can streamline traditionally time-consuming workflows while accelerating alpha generation. Additionally, the creation of internal copilots that enable faster synthesis of this data will drive productivity gains.

3. Organizational efficiencies and cost reduction in insurance

Unstructured data underpins the insurance industry — across contracts, claim notes, policy documents and more. Gen AI tools empower insurers to reduce organizational costs while increasing efficiencies around claims, underwriting and customer service. LLMs can help with everything from categorizing incoming emails and prioritizing calls to assessing risk and retrieving policy details from historical records.

4. Reduced exposure to risk and improved regulatory reporting

Risk management and compliance teams can leverage generative AI to simulate adverse market conditions and potential future risks, allowing them to prepare for worst-case scenarios. They can also use gen AI to generate synthetic data to use for stress-testing financial models and identifying new patterns of fraud.

5. Democratized data science capabilities

Gen AI tools allow non-data scientists and non-technical employees to take advantage of the power of data science by generating code with the use of plain natural language. This means that teams can ideate, develop and build new capabilities faster.

PREPARING FOR
THE AI FUTURE

In the last fiscal year in the **Streamlit** developer community, we saw 20,076 unique developers work on 33,143 LLM-powered apps (including apps still in development).

That means that the future filled with the power of AI is here. It may not be evenly distributed yet, but it's here.

Our research strongly suggests 2024 will be the year that generative AI becomes a crucial component of the enterprise technology stack, and that gen AI applications will spur better data-driven decision-making. It's also clear that unified data management and governance will be critical to making those efforts a success, both for financial services companies and the greater business community.

Today's forward-thinking organizations are creating more complex LLM applications, making AI more available across the enterprise, and seeing the benefits of a unified data platform. Thanks to the power of the Snowflake Data Cloud, the hype surrounding AI is beginning to transform into real business benefits.

Learn more about how to build for success and prepare for the AI future with the Financial Services Data Cloud.

APPENDIX: METHODOLOGY

The Snowflake Data Trends 2024: Financial Services report is generated from fully aggregated, anonymized data detailing usage of the Snowflake Data Cloud and its integrated features and tools. In this report, we examine patterns and trends in data and AI adoption across Snowflake's financial services accounts. These trends provide insight into the state of data and AI, including which technologies are the fastest growing within the financial services space. Note that usage attributable to internal consumption, if any, has been removed and is not reflected in any of the metrics contained herein. The accounts and usage reflected in this report include both longtime Snowflake users and others who only recently joined the Data Cloud.

Except where noted in the text, the data in this report compares monthly averages from January 2024 (represented as "this year") to averages in January 2023 ("last year"). When compared, this is depicted as "year over year" growth to align with Snowflake's fiscal year-end, though the figures themselves are only representative of January figures to calculate growth.

When possible, we have provided these year-over-year comparisons to showcase growth trends over time. Where data was drawn from Snowflake features that became publicly available after the start of the fiscal year, data was collected and compared as of the first full month after which the feature became available in public preview, and that date is noted in the text. Notably, growth figures for features moving into public preview are expected to be considerably higher, as private previews are limited in scope and necessarily restricted to select Snowflake customers.

ABOUT SNOWFLAKE

Snowflake enables every organization to mobilize their data with Snowflake's Data Cloud. Customers use the Data Cloud to unite siloed data, discover and securely share data, and execute diverse artificial intelligence (AI) / machine learning (ML) and analytic workloads. Wherever data or users live, Snowflake delivers a single data experience that spans multiple clouds and geographies. Thousands of customers across many industries, including 691 of the 2023 Forbes Global 2000 (G2K) as of January 31, 2024, use the Snowflake Data Cloud to power their businesses.

Learn more at snowflake.com