



HOW HEALTHCARE AND LIFE SCIENCES COMPANIES CAN LEVERAGE THIRD-PARTY DATA FOR ANALYTICS

Enrich your data and unlock new insights faster with live, governed access to external data



CHAMPION
GUIDES

EBOOK

TABLE OF CONTENTS

- 3** Executive summary
- 4** Why frictionless access to third-party data can improve patient, clinical, and business outcomes
- 5** Two ways to unlock hidden value in your data
- 6** The challenges of sourcing third-party data
- 7** The Healthcare and Life Sciences Data Cloud solves traditional data sharing issues
- 8** Snowflake Marketplace
- 9** Snowflake customer use cases
- 10** Become an empowered data consumer today
- 11** About Snowflake



EXECUTIVE SUMMARY

Third-party data, which is data that comes from sources external to an organization, enables healthcare and life sciences organizations to improve clinical and patient outcomes as well as the effectiveness of go-to-market and marketing strategies. External data empowers teams to make better, data-driven business and clinical decisions, especially when it's integrated with first-party data or data a company collects directly from its customers.

However, traditional methods for sourcing third-party data are inefficient and insecure. Traditional data marketplaces may not scale and legacy technologies for transferring data (FTP, APIs) can require extensive engineering work before the data can be used. This situation can result in delays, stale, and poor data analysis.

In this ebook, you will learn how to:

- Access live third-party data without any ETL, making the data immediately available for analysis or to merge with your own data
- Easily discover third-party data sets, such as anonymized prescription data, medical sales data, and demographic data that best fit your business needs
- Use enrichment services to improve the quality of first-party data by securely sharing slices of your data with providers

WHY FRICTIONLESS ACCESS TO THIRD-PARTY DATA CAN IMPROVE PATIENT, CLINICAL, AND BUSINESS OUTCOMES

Companies in the healthcare and life sciences industries are becoming more data-driven, which has a number of implications. Pharmaceutical companies are optimizing drug development and commercialization through borderless data access, while payers and providers are improving patient outcomes through a more secure and seamless exchange of patient information. Meanwhile, distribution partners are sharing real-time insights, which helps to minimize supply chain disruptions.

Although healthcare and life sciences companies have their own rich insights on patients, prescriptions, and outcomes, they're increasingly leveraging third-party data to obtain additional insights and enhance targeting for the purposes of improving clinical trials and marketing ROI. The problem is accessing that data in a timely, cost-effective way that is compliant with industry data privacy and governance regulations such as the U.S. Health Insurance Portability and Accountability Act (HIPAA) and the EU's General Data Protection Regulation (GDPR). Healthcare companies often build and maintain dozens, if not hundreds, of data pipelines and deal with a wide range of data formats, which can result in their data scientists spending much of their time on relatively low-value business data manipulation and conversion tasks. Complex ingestion processes can also result in reductions in data quality and a high error rate

SOURCES OF THIRD-PARTY DATA

Generated by individuals, businesses, medical devices, IoT sensors, and more, third-party data originates from a variety of sources and exists in a wide range of categories, including but not limited to the following:

- **Consumers (demographics, consumer sentiment, psychographics)**
- **Health (public health, genomic annotation, prescriptions, life expectancy)**
- **Point of sale (online, in-store)**
- **Mobility (location, in-store foot traffic)**
- **Online behavior (searches, social media, app usage, web traffic)**
- **Media/advertising (consumption, measurement, viewership)**
- **Events (satellite, weather, event detection)**

By combining these third-party data sets with their own data, healthcare and life sciences companies can improve the accuracy of clinical trials, optimize patient care and follow-up, and build a 360-degree view of customers to personalize content and offers and improve ad targeting.



TWO WAYS TO UNLOCK HIDDEN VALUE IN YOUR DATA

The benefits of using third-party data from a data provider are clear but companies in the healthcare and life sciences industries can extract even more value from external data when they integrate it with first-party data and analytics.

Below are two examples of how third-party data, when combined with first-party data sets, can optimize clinical trials and improve the performance of marketing campaigns.

IMPROVE THE EFFICIENCY OF CLINICAL TRIALS

By leveraging genomic annotation data, real-world data from platforms such as IQVIA and Compile, and other third-party data sets, life sciences companies can realize more efficient clinical trials and accelerate standard timeframes for bringing a new drug to market.

Third party data is also utilized to accelerate the identification and recruitment for clinical trial participants. Life science organizations, both biopharmaceutical and medical device, are acquiring electronic health records (EHR), medical claims, pharmacy, and real-world data to proactively identify and recruit patients for participation in clinical trials. Many mature organizations have also created or partnered with other organizations to develop a patient cohort builder that can be a self-service application across all therapeutic areas for accelerated identification and predictor of patient adherence.

PERSONALIZE CONTENT AND ADS TO IMPROVE MARKETING ROI

By leveraging third-party demographic data, life event data, purchase data, and browsing data, healthcare and life science companies can enrich the customer information that is already in their databases and obtain a true patient/customer 360. Refining their segmentation and targeting strategies to increase the likelihood of patients, members, and customers receiving content and offers that are truly relevant to them makes marketing spend more efficient.

For example, payers might advance value-based health management initiatives for at-risk member populations by ensuring that the right audience receives information about available care plans and services. This will assist in the management of chronic diseases, such as diabetes and heart disease, by ensuring that patients receive relevant information about carrier-sponsored wellness programs and benefit offerings at every stage of their health journey.

THE CHALLENGES OF SOURCING THIRD-PARTY DATA

While the reasons to source third-party data are sound, the legacy practices for doing so are inefficient. Two major obstacles stand in the way for many healthcare and life science companies:

- **Expensive:** Traditional methods of accessing external data require extensive time and monetary resources
- **Unsecure:** Legacy file sharing methodologies (FTP/APIs) are error-prone and not secure

Getting data from traditional data marketplaces can also be overwhelming. How do you choose from the multitude of data marketplaces? How do you get started? How do you determine which vendors are reliable, what data will be most useful, and what value each data set really holds?

These questions point to the cost, time, and effort required to find and select the best-suited third-party data. It comes down to a problem of scalability. There's no efficient process for contacting each vendor, evaluating its data, and acquiring the data.

Organizations spend enormous amounts of time and money on teams to perform this arduous task. Some organizations let their data scientists muddle through the data acquisition process, but this takes expensive resources away from the task they were hired to do: build data models. Other organizations use data aggregators and brokers to facilitate data marketplace transactions. This solution has advantages but doesn't address the challenges presented by reliance on traditional file sharing methodologies.

5 BIGGEST CHALLENGES WITH TRADITIONAL DATA SHARING METHODS:

- **Stale, questionable data and poor analytics results**
- **Potential compliance and security issues**
- **Time-consuming and costly**
- **Impossible to securely govern**
- **Prone to human error**

Vendors continue to employ FTP, APIs, and other file downloading techniques to transfer data, which often requires them to copy files and perform intensive engineering work to extract, transform, and load (ETL) data. Even APIs become a burden when developers need to maintain and troubleshoot multiple APIs from separate data providers. Every API has different security and authentication methods, which adds work for an organization's security teams and poses potential risks. APIs are also inefficient for receiving large volumes of data.

But the biggest challenge with these data sharing methods is that they result in stale data copies that are nearly impossible to securely govern. Manual transfer efforts are prone to human error, leaving organizations exposed to potential security and compliance issues. Worst of all, they end up with questionable data that results in poor analytics.

Taken together, these challenges raise this question: How can you use external data at scale without wasting time, money, and resources and without





THE HEALTHCARE AND LIFE SCIENCES DATA CLOUD HELPS SOLVE TRADITIONAL DATA SHARING ISSUES

If organizations could minimize time spent on building and managing ad hoc methods of data sharing, they could allocate more resources to extracting valuable insights from that data. But the numerous barriers presented by traditional data marketplaces and legacy data sharing practices must be eliminated so organizations are empowered to discover and evaluate data sources with ease and combine external data with internal data for rapid analysis.

Snowflake's Healthcare and Life Sciences Data Cloud represents the modern answer for using data—of all types. With its multi-cluster shared data architecture, Snowflake's platform centralizes data in a single, secure location in the Data Cloud; the network connects Snowflake customers, partners, data providers, and data service providers across public cloud providers and regions. The result: traditional data sharing barriers are removed and data silos are eliminated. Organizations immediately benefit from access to secure and governed data, which can be shared within and between organizations.

Snowflake's global platform centralizes data in a single, secure location in the Data Cloud, making it uniquely designed to connect businesses around the world and power seamless data collaboration and many different workloads.

That's because Snowflake is built using **Snowgrid**, a technology that connects regions and clouds around the globe and enables secure and governed data sharing.

With Snowgrid, you can take advantage of **Snowflake Collaboration** features to seamlessly and securely share and access live, ready-to-query data across clouds and regions without any ETL or APIs. Anyone granted access to a data set simply references the data in a controlled and secure manner without gaining physical custody of the data. This means that data access is revocable, enabling you to better comply with regulations such as GDPR. And because any changes made to the data are done to a single version, data remains up to date for all data consumers who have access, with minimal latency or concurrency issues.

With first-, second-, and third-party data unified in the Data Cloud, organizations can combine shared data instantly with existing data for faster analysis. Data is available in a ready-to-query format without replication, transformation, or processing, so data analysis delays will become a problem of the past.

And, with its cloud-agnostic architecture, Snowflake enables organizations to have seamless and immediate access to shared data, regardless of cloud infrastructure, cloud provider (AWS, Azure, or Google Cloud Platform), or supported region.

SNOWFLAKE MARKETPLACE

Snowflake Marketplace is a modern approach to the data marketplace. Data consumers can access live, up-to-date, and ready-to-query third-party data, data services, and [Snowflake Native Apps](#) from multiple providers in one place. Rather than wasting time hunting for vendors and downloading stale data, consumers can use Snowflake Marketplace to easily discover, evaluate, and buy live external data that is secure, enabling compliance with regulatory requirements.

With the burden of data transformation removed, integrating external data with your existing data becomes fast and seamless. Data analysis can begin right away.

Improve your own data quality with enrichment services that don't require the traditional, time-consuming steps of data copying and transformation. Through Snowflake Marketplace, you can share slices of your data with the data provider for enrichment and augmentation, and get your enriched data securely shared back directly into your Snowflake account.

Whether you use external data to augment data sets for improved business analytics or to help train AI/machine learning (ML) models for data science, the benefits of Snowflake Marketplace are clear.

- **Ease of discovery:** Benefit from a single point of access for a wide variety of data sets that can be queried, joined with internal data, used in data modeling, or added to BI tools—all with speed and ease.
- **Live, up-to-date data:** With no manual intervention or scheduling required, all updates made by the third-party data provider are immediately reflected in your data sets.
- **Reduced costs:** Eliminate unnecessary data analytics expenses around data loading, transformation, and API integration and management. Because there's no data movement (only data access), you also don't pay storage costs for third-party data.
- **Personalization:** Request personalized, secure data feeds from data providers that are customized to your specific data needs.
- **Enriched internal data:** Use enrichment services to improve the data quality of your own data by securely sharing slices of your data with providers.
- **Global access:** Receive fast access to third-party data on multiple major cloud providers.

**SNOWFLAKE MARKETPLACE
CONNECTS YOU TO OVER 430
PROVIDERS, OFFERING MORE THAN
1,800 LIVE, READY-TO-USE DATA SETS,
DATA SERVICES, AND SNOWFLAKE
NATIVE APPS.***

HCLS DATA SETS, including:

- IQVIA–Data-as-a-Service
- Invitae–Variant Interpretations
- PRECISIONxtract–Proprietary Payer Data Master
- GlobalData–Pharma Drug Sales & Forecast Database
- Innovaccor–Data Activation Platform
- SameSky Health–Consumer Demographic Analysis
- Syntegra–Patient-Level Clinical Data (Inpatient and Outpatient Hospital), and more

*As of April 30, 2023.

SNOWFLAKE CUSTOMER USE CASES



U.S. Biotechnology company

EXPEDITE DRUG COMMERCIALIZATION

Pfizer has used Snowflake to help digitize drug discovery and manufacturing, enabling the company to improve patient outcomes, reimagine clinical trials through robotics and automation, and perform predictive analytics for diagnosis and supply chain tracking.

Pfizer has also been able to drive marketing and sales performance with faster insights and improve safety and risk management. With the help of Snowflake, Pfizer's IT team spends less time wrangling data and more time analyzing it to help provide better treatment solutions for patients. Some of the team's data pipeline jobs have dropped from 12 hours to five minutes. Snowflake also enables Pfizer to collaborate more easily and efficiently with other organizations, including syndicated data partners and regulatory agencies. Additionally, Pfizer has built a global interaction repository—a core model that stores all of its digital interactions—that analysts can tap into using the same analytical workflows and techniques to create business value and improve patient outcomes.



Switzerland Pharmaceuticals company

ACCELERATE DRUG DISCOVERY AND DEVELOPMENT

Novartis, a leading global medicines company, uses innovative science and digital technologies to create transformative treatments and products for nearly 800 million people around the world. Novartis had a wide spectrum of data types and data management systems that were not standardized, interoperable, or scalable. It also must abide by strict industry regulations on data privacy and governance. As a result, it could take up to six months to get insights from data. Now, with the Healthcare & Life Sciences Data Cloud, Novartis achieves interoperability across multitudes of vendors, systems, and data sources. The Healthcare & Life Sciences Data Cloud has helped Novartis empower innovation and improve effectiveness for analytics and data science teams by removing data bottlenecks and barriers to insight, and enabled a global data strategy by leveraging multi-cloud capabilities.



U.S. Healthcare Insurance company

DRIVE PATIENT CENTRICITY AND PERSONALIZATION

Elevance Health serves more than 110 million people through its affiliated companies, including approximately 43 million members within its own family of health plans. The company has pioneered a “whole-person” approach to healthcare in which it considers multiple factors—including social drivers of health—that influences its members' well-being. But leveraging all of the available data to better serve patients and improve patient health outcomes was a challenge. Legacy systems were on-premises, making scalability difficult and creating significant delays in getting important patient insights. Using Snowflake's Data Cloud, Elevance Health can now assess an individual's condition or predict future health needs, drawing on a wide range of data, including historical, social, and environmental information. This includes prior insurance claims, electronic medical records, and hospital visits, as well as biometrics and sensor data. Elevance Health is also using the Data Cloud for machine learning and artificial intelligence workloads.

Learn more about the success healthcare and life sciences organizations are having with the Snowflake Healthcare & Life Sciences Data Cloud.

BECOME AN EMPOWERED DATA CONSUMER TODAY

Snowflake Marketplace gives you the ability to securely access and quickly combine data from third-party sources, helping to improve the rigor of clinical trials and boost campaign performance when new drugs or medical devices are taken to market.

Discover for yourself the difference Snowflake Marketplace makes. Go to snowflake.com/marketplace/ to browse listings and sign up for a **free trial**. There, you can find genomic annotation data, anonymized prescription and medical claims data, COVID-19 data, health expenditure data, and more, and can start querying these data sets immediately.

There's only one question remaining:
Are you ready to extract more value
from your first-party data?





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, power data applications, and execute diverse AI/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 590 of the 2022 Forbes Global 2000 (G2K) as of April 30, 2023, use Snowflake Data Cloud to power their businesses.

Learn more at [snowflake.com](https://www.snowflake.com)



© 2023 Snowflake Inc. All rights reserved. Snowflake, the Snowflake logo, and all other Snowflake product, feature and service names mentioned herein are registered trademarks or trademarks of Snowflake Inc. in the United States and other countries. All other brand names or logos mentioned or used herein are for identification purposes only and may be the trademarks of their respective holder(s). Snowflake may not be associated with, or be sponsored or endorsed by, any such holder(s).