5 STRATEGIES TO IMPROVE SECURE DATA COLLABORATION

A comprehensive guide to creating your own data exchange
# TABLE OF CONTENTS

3  The Power of Data Sharing
4  Secure Data Sharing Changes Everything
6  Create Your Own Data Exchange
7  Which Data Exchange Strategy Is Best for Your Organization?
8  Five Strategies for Data Collaboration
9  Strategy #1: Intra-Company
11 Strategy #2: Cross-Company
13 Strategy #3: Data Sourcing
15 Strategy #4: Data Distribution
17 Strategy #5: Hybrid
18 Are You Ready for a Data Exchange?
19 About Snowflake
Today, data is an imperative for organizations across all fields, regardless of their size or purpose. Strategically, businesses rely on data-driven insights into industries, operations, and customers to remain competitive, improve performance, and deliver strong user experiences. For industries such as healthcare, transportation, aviation, construction, and banking, real-time data analysis is becoming a necessity to ensure the health and safety of people, especially as automation and reliance on artificial intelligence increase.

In 2025, IDC predicts that 175 zettabytes (ZB) of global data will be generated, which represents 10 times the amount of data created in 2016. The need to share data takes on a new importance in light of its criticality in the global economy. Organizations want to break down data silos and derive insights faster by enabling intra-company and cross-company sharing with partners, vendors, suppliers, and customers. While some see data sharing as an opportunity to monetize their data, others want to consume data from more sources to enrich their own analysis.

Unfortunately, sharing data represents a real challenge for organizations that rely on traditional methods. Any time data sharing includes emailing spreadsheets or conducting batch processes such as FTP and ETL, the result is stale data copies that waste valuable time, money, and resources.

Data-driven organizations now require a better way to share data, one that does not require copying or even moving any data and eliminates the need to transform data so it can be used.
SECURE DATA SHARING CHANGES EVERYTHING

Many organizations struggle to share files internally across departments and externally with partners, vendors, suppliers, and customers. Manual methods are often used, including emailing spreadsheets or executing batch processes that require extracting, copying, moving, and reloading data.

These traditional data sharing methods are notorious for their lack of scalability and security, and their error-prone and costly processes are time-consuming for data teams. Worst of all, after investing significant time and resources to prepare the data for consumption, it has become stale. Those who try to use a basic cloud storage service quickly discover it’s equally inefficient and lacks the ability to query data directly.

Thankfully, data sharing is a natural function of a modern platform such as Snowflake’s Data Cloud. Snowflake’s architecture separates compute, storage, and services to allow unlimited scaling and concurrency. This architecture enables Snowflake’s secure data sharing to do two things:

• Never make copies of data
• Never move data

Instead, Snowflake’s secure data sharing is about accessing data. Live data is always shared from its original location, and those who are granted access simply reference the data in a controlled and secure manner.

This is important: Data doesn’t move. Because there are no data copies or any transmission of data, a single version of the data exists, and that data remains current and centralized. Data access is provided at a granular level, such as access to specific databases, tables, or rows of data, and any real-time changes to data are available instantly to data consumers.

And we can promise you that no one will miss ETL or FTP processing. With secure data sharing, data is available immediately in ready-to-query format without any transformation, which enables shared data to be combined with existing data for faster analysis and stronger insights. The data consumer experience improves dramatically.

Snowflake is truly cross-cloud, which means secure data sharing works across cloud providers and regions, ensuring seamless and immediate access. Secure data sharing also works for sharing data with companies that don’t have a Snowflake account, which makes it a global and inclusive feature.

THE PITFALLS OF TRADITIONAL DATA SHARING METHODS

Traditional data sharing and distribution methods often use technology such as FTP, cloud buckets, or APIs. These methods have the following disadvantages:

• Storage costs for both parties
• ETL costs and effort for both parties
• Security vulnerabilities
• Service and support costs
• Latency and potential errors leading to poor customer experience

If you are just starting out in your data exchange journey, you might be tempted to develop APIs as a way of sharing data. Although APIs are a great way to connect different systems and automate processes, they have a series of additional challenges when they are used for data exchange, including:

• Requiring in-house expertise to develop and maintain them
• Requiring recurring effort and costs to develop and maintain them
• Limiting the volume of data available for access
• Requiring data consumers to learn to use the API
• Limiting the types of questions the data consumer can ask against the data
• Causing performance and quality issues that are difficult to resolve
Figure 1: Snowflake facilitates secure data sharing across cloud providers and regions.
CREATE YOUR OWN DATA EXCHANGE

Imagine all barriers to sharing and monetizing your data disappear forever. Think of the multitude of ways you could transform your organization by sharing and receiving data in a curated, governed, and secure manner, both internally and externally, and at scale.

With secure data sharing, this mental exercise is now an actuality and illustrates why Snowflake has enabled organizations to create their own data exchanges, built on top of secure data sharing, as a new way to frictionlessly share data.

In essence, data exchanges offer a secure location where consumers can easily discover data that has been published by providers. Consumers then get access to this data through Snowflake’s secure data sharing capability. Data exchanges enable:

- **Discovery** of trusted data sets where authorized users can easily find live, ready-to-query data that’s been made available to them
- **Control** to invite and manage who can be data providers and data consumers
- **Management** of listings and the creation of different packages for data consumers
- **Metrics** to track and monitor data usage throughout the exchange
- **Monetization** by opening up new ways to provide data services and monetize data

Data exchanges demolish data silos. They remove all the friction from collaboration between departments within a company and open up new avenues for data sharing with key partners, customers, vendors, and suppliers. When you create your own data exchange on Snowflake, you have complete control over who provides, views, accesses, and consumes the data.

In addition, data exchanges provide an easy and secure way to become a data provider and monetize your data for new revenue streams. The reverse is also true: Data exchanges enable you to act as a data consumer, so you can browse listings within different data categories and acquire third-party data to combine with your own data for new insights. Best of all, shared data from other sources doesn’t consume storage in your account. When you access and use that data, you’re simply pointing to the original account’s data.

Just like secure data sharing, data exchanges work cross-cloud, across regions, and enable you to share data with and access data from companies that don’t have Snowflake accounts.

POWERED BY SECURE DATA SHARING, DATA EXCHANGES ENABLE:

- **Scale**: Share data with dozens, hundreds, or even thousands of other entities securely and cost-effectively.
- **Governance**: Create and manage data in a centralized location where authorized users can access it without increasing cost and concurrency issues.
- **Control**: Manage dozens, hundreds, or thousands of different entities accessing the same data sets without losing control.
- **Visibility**: Track what data is being accessed, who is accessing it, how often, and for how long.
- **Discovery**: Ensure that employees, customers, business partners, and others in your ecosystem can easily discover approved and curated data sets they can use in their own operations, without requiring data transformation.
- **Monetization**: Make certain data sets available to customers who might be interested in paying for access or data services, and create a monetization strategy that enables your organization to get more value out of your data.
WHICH DATA EXCHANGE STRATEGY IS BEST FOR YOUR ORGANIZATION?

Data exchanges redefine how companies acquire, leverage, and monetize data to power business and improve their ecosystems. To discover which data exchange strategy is a best fit for your organization, consider the following:

1. Are you focused on breaking down internal data silos, controlling and governing access to data, and improving overall operational efficiency for faster time to market?

2. Would you like to improve collaboration with external partners, vendors, customers, and suppliers in the hope of opening up new monetization opportunities and improving time to market?

3. Are you looking for new sources of data to complement your current data in order to improve your offerings through richer analytics and more informed decision-making?

4. Do you wish to share and monetize your data at scale to create new revenue streams and increase profits?

5. Did you answer "yes" to two or more of the questions above?

These questions correspond to five strategies for data collaboration:

1. Intra-company strategy
2. Cross-company strategy
3. Data sourcing strategy
4. Data distribution strategy
5. Hybrid strategy
FIVE STRATEGIES FOR DATA COLLABORATION
It’s not unusual to struggle with data sharing across business units, departments, locations, and subsidiaries, especially if you’re a national, multinational, or global organization. More often than not, a cohesive data strategy was missing at the outset. Urgency took the place of cohesion, and teams and departments adopted various data storage solutions to meet their pressing needs.

These one-off decisions combined to form the data silos of today. As a result, data-driven insights and business opportunities are often limited in value and scope due to incompatible systems and the inability to share real-time data with colleagues.

Here are some tell-tale signs that your organization is facing challenges with siloed data:

- Do employees across different functions, business units, or regions submit constant requests for access to data sets outside their group or location?
- Are you seeing escalating costs to fulfill these data requests, as well as a drain on data scientists and other departments that manage requests?
- Are there specific data sets within the company that should be accessible and leveraged by more employees and groups?

If you want to deliver stronger business productivity, better internal collaboration, and more data-driven insights, it’s time to maintain a central location for approved and secure data.

<table>
<thead>
<tr>
<th>Why create a data exchange?</th>
<th>Break down internal data silos to improve operational efficiency.</th>
</tr>
</thead>
<tbody>
<tr>
<td>How it works</td>
<td>• Create your data exchange and invite different business units to be members for data publishing and access.</td>
</tr>
<tr>
<td></td>
<td>• Set up a publishing workflow to ensure data follows data governance guidelines.</td>
</tr>
<tr>
<td></td>
<td>• Create policies regarding the listing of data sets.</td>
</tr>
<tr>
<td></td>
<td>• Monitor data usage and improve data collaboration.</td>
</tr>
</tbody>
</table>

**Snowflake customer example**

A large financial services provider wanted to grant data scientists from different business units access to and visibility into the approved data sets of other departments while maintaining data security and governance. With a data exchange, the company’s data scientists can now source company-approved data from a central location with speed and agility, which delivers stronger analysis across the entire organization.
INTRA-COMPANY STRATEGY
Value: Break down internal data silos to improve operational efficiency

Your Data Exchange

Figure 2: Example of intra-company data exchange.
Organizations today are operating within a global economy that demands a new level of collaboration with external partners, suppliers, vendors, and customers. However, fear around controlling data and not wanting copies to exist outside of the organization often lead to locking down proprietary data and not benefiting from cross-company collaboration.

Companies that will continue to be competitive are those that recognize the value in sharing information internally and externally to spur growth and productivity. The act of data sharing strengthens collaboration, accelerates new insights, and improves business relationships.

A data exchange can improve cross-company collaboration through enabling internal employees to make data sets accessible to external customers, suppliers, vendors, and partners, while maintaining control over their own data, which data sets are shared, who has access, and who can publish data sets.

Ask these questions to determine whether a cross-company data exchange strategy is right for you:

- Do you routinely share data, either via FTP, emails, cloud storage, or other means, with companies outside your organization?
- Do customers and business partners request access to your data, which is outside the typical service you provide today?
- Is receiving data from vendors or business partners taking time and resources away from other departments and considered a “pain”?
- Are you sourcing third-party data from different data vendors, and would streamlining and centralizing this process be beneficial to the company?
- As you map out your business relationships with customers, vendors, partners, and other entities, is there a need for better data flow that could improve supply chain efficiency and business relationships?

**Why create a data exchange?**

Exchange data with other organizations to improve collaboration and open up monetization opportunities.

**How it works**

- Create your data exchange and invite other organizations to publish data, access data, or do both, regardless of their cloud provider or region.
- Establish data quality guidelines and set up a data publishing workflow.
- Create multiple data categories to enable easy data discovery.
- Monitor data usage and track data access.

**Snowflake customer example**

A multinational technology company built a data exchange to enable frictionless data sharing with its parts suppliers. This allows suppliers to get immediate access to critical product quality data and perform root-cause analysis to discover defects and suggest remediations. The resulting data is shared back with the company and across business units, which then assess the impact on production schedules, sales targets, and more. By improving data collaboration across its supply chain, the technology company has reduced costs for data sharing and data analytics as well as improved its operational efficiency.
CROSS-COMPANY STRATEGY

Value: Exchange data with other organizations to improve collaboration and open up monetization opportunities

Your Data Exchange

Figure 3: Example of cross-company data exchange.
Companies in industries that rely on external data for competitive advantage are always looking for better ways to get unique data sets and generate new insights. Financial services organizations such as hedge funds rely on alternative data which, when combined with the standard market data, can give them new perspectives on why certain businesses are doing well and what particular trends could adversely affect their portfolio. However, the sheer number of different data sources, the different file formats and delivery methods used by data vendors, and the data transformation efforts required by the data scientists create a bottleneck that limits data volume and how frequently the data can be sourced, evaluated, acquired, and finally put to use, not to mention the costs associated.

For these companies, having a better way to source external data at scale in a governed way is critical for their ongoing success. By creating their own data exchange, they can invite data vendors to publish their myriad data sets to a central portal where data scientists can easily and quickly evaluate the data and decide which data to acquire. Once that decision is made, because the data from the vendors is already on that company’s data exchange, the data scientists have immediate access to the data directly from their data warehouse.

By creating your own data exchange, not only will your teams access live, ready-to-query data that is preapproved, but you can track data usage and have complete visibility into which data sets and providers are used the most. While ensuring security, you will reduce time to insight and improve collaboration across a broader spectrum of data.

Organizations looking to deploy a data exchange to primarily streamline data sourcing are typically answering yes to the following questions:

- Is your company relying on third-party data as part of key competitive differentiation?
- Are you sourcing data from a large number of data sellers?
- Would streamlining the data sourcing, evaluation, and acquisition process give your company a significant edge and enable the company to improve data analytics?

<table>
<thead>
<tr>
<th>Why create a data exchange?</th>
<th>Improve external data sourcing and reduce costs for data analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How it works</strong></td>
<td>• Identify all the data sellers you work with today and invite them to become members in your own data exchange.</td>
</tr>
<tr>
<td></td>
<td>• Set up a data approval workflow to ensure data is reviewed before it is made accessible to your data scientists or data buyers.</td>
</tr>
<tr>
<td></td>
<td>• Provide metadata and example queries in data listings to help data analysts evaluate the different data sets faster.</td>
</tr>
<tr>
<td></td>
<td>• Use ready-to-query data sets to enable data analysts to make faster decisions.</td>
</tr>
</tbody>
</table>

**Snowflake customer example**

A hedge fund set up a private exchange for its team of data scientists. Every year, alternative data is sourced from hundreds of different vendors in order to understand market trends and determine where money should be invested. In order to reduce data sourcing costs and streamline the data acquisition process, the firm’s data exchange now provides a single location for vendors to upload all of their data, ensuring data buyers and data scientists have a central place from which to evaluate data, and ensuring data is in a ready-to-query format that speeds up the data evaluation and acquisition process. This saves the company the cost and effort previously required to access and transform data and gives them a better way to evaluate data and make data-driven decisions.
DATA SOURCING STRATEGY
Value: Improve external data sourcing and reduce costs for data analysis

Your Data Exchange

Figure 4: Example of a data sourcing data exchange.
As a source of data for other organizations, your company has likely struggled to deliver data without workarounds. Perhaps you’ve used APIs, copied files using FTP, or tried other traditional data sharing methods in order to give your data customers access, while realizing fully that they are receiving only a snapshot in time of the data. Even cloud buckets have fallen short due to friction and lack of scalability. Isn’t it frustrating not to be able to fulfill customers’ requests for direct data access?

All that changes with a data exchange. By creating and hosting your own walled garden of shared data, you can provide customers with an online data portal where live data is accessed securely without the need for transformation.

In addition to direct data access, data feeds can be personalized for better customer experiences and increased customer retention. By monitoring and tracking data usage, you can ascertain which data sets are most popular and improve your data products and monetization strategies. With the removal of API maintenance or the use of FTP, you also see an overall reduction in costs and improved margins.

To determine whether a data exchange could improve the way you provide data access to outside entities, such as customers, ask yourself the following:

- Are you getting requests from customers or other third-parties for raw or granular data access?
- Would giving those organizations easy and secure access to data help your company to strengthen business relationships, improve customer experience, or open new business opportunities?
- Is data monetization an important part (at least in concept) of your company’s business strategy?

<table>
<thead>
<tr>
<th>Why create a data exchange?</th>
<th>Share data with other organizations, at scale, and potentially monetize your data assets.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Set up your data exchange and invite customers or other organizations to become data consumers.</td>
</tr>
<tr>
<td></td>
<td>• Create different data categories for easy data discovery.</td>
</tr>
<tr>
<td>How it works</td>
<td>• Publish metadata and usage examples alongside data sets to showcase the value of the data that is offered.</td>
</tr>
<tr>
<td></td>
<td>• Publish live data so other organizations accessing the data will always have the most up-to-date data sets.</td>
</tr>
<tr>
<td></td>
<td>• Monitor data usage and find out which are the most sought-after data sets.</td>
</tr>
<tr>
<td></td>
<td>• Create new data products and monetize them.</td>
</tr>
<tr>
<td>Snowflake customer example</td>
<td>An educational technology company created a secure data platform to service its customers. This centralized location provides all the data the organization wants to share externally with customers and enables seamless access to the right data for each customer type.</td>
</tr>
</tbody>
</table>
DATA DISTRIBUTION STRATEGY

Value: Share data with other organizations, at scale, and potentially monetize your data assets

Your Data Exchange

Figure 5: Example of a data distribution exchange.
#5 HYBRID STRATEGY

If you find yourself enticed by more than one type of data collaboration, chances are that your organization would benefit from a hybrid solution. After all, there’s no reason to limit collaboration to only one scenario.

For example, you can promote intra-company and cross-company data collaboration at the same time as you set up a data exchange account to source the best external data sources possible. Many Snowflake customers fall into this hybrid category. When you think broadly, there are no limits to the ways in which you can publish, access, and consume data in this new world of data exchanges.
ARE YOU READY FOR A DATA EXCHANGE?

Every organization benefits from real-time and seamless access to data. Only a narrow view of the world can be examined when data is confined to one organization—or worse, to individual business units. Comprehensive, impactful data analysis requires a breadth and depth of information that few, if any, companies possess on their own.

To match the pace of business and thrive in this global economy, the goal should be to foster transparent, data-driven relationships within and across organizations.

Is your company ready to capitalize on data sharing and create data exchanges? These guiding questions are designed to help you think through the steps necessary to bring data sharing to your organization and involve others in that process.

What data does your company source today? Here are some examples:
- Credit data for the accounting department
- Demographic and technographic data for marketing
- Weather and traffic data for production
- Alternative data for data scientists

What data would you like to source but you can’t because:
- It’s too expensive
- It’s too complex
- You do not have the resources
- You’re not sure where to get it from

Of the business relationships your company has today:
- Are customers requesting access to data?
- Could better data pipelines help with supplier relationships?
- What business partners would benefit from your data (and vice versa)?
- What would happen if access to data was made easier, simpler, and less expensive?

These questions are the tip of the iceberg, but we hope they spark the necessary dialogue for introducing data exchanges to your organization. Then the only query that will remain is, what’s your ideal scenario for data collaboration?
ABOUT SNOWFLAKE

Snowflake delivers the Data Cloud—a global network where thousands of organizations mobilize data with near-unlimited scale, concurrency, and performance. Inside the Data Cloud, organizations unite their siloed data, easily discover and securely share governed data, and execute diverse analytic workloads. Wherever data or users live, Snowflake delivers a single and seamless experience across multiple public clouds. Snowflake’s platform is the engine that powers and provides access to the Data Cloud, creating a solution for data warehousing, data lakes, data engineering, data science, data application development, and data sharing. Join Snowflake customers, partners, and data providers already taking their businesses to new frontiers in the Data Cloud. snowflake.com

© 2021 Snowflake Inc. All rights reserved.