

EBOOK

The Five Stages of Building an Agile Data Infrastructure for Enterprise Reporting

Contents

The evolution of enterprise reporting	3
The data problem	4
What is an agile data infrastructure?	5
What are the challenges organizations face in enterprise reporting?	6
How to build an agile data infrastructure for enterprise reporting in five stages	7
Three key enterprise reporting use cases	13
What advantages does an agile data infrastructure provide for enterprise reporting?	18



The evolution of enterprise reporting

Enterprise reporting was once a relatively simple task. Teams would take data from their ERP, finance, and other internal systems to file a set number of reports at scheduled intervals. While that data was static and almost immediately outdated, it was usually sufficient.

But enterprise reporting has evolved. The days of only having to report on a predictable monthly or quarterly cycle are over. Expectations and stakeholder demands have changed significantly thanks to three main drivers:

- **Customers, suppliers and employees** want organizational transparency.
- **Investors** want accurate and trusted data.
- **Regulators** are constantly increasing the reporting burden.

What this translates to is **continual reporting on multiple fronts**, from financial reporting to regulatory compliance, to environmental, social, and governance (ESG) disclosures. This isn't going to change. Reporting will continue to increase in depth, breadth, complexity, and frequency.

Satisfying reporting obligations that differ by region and industry is difficult, but non-compliance can mean severe fines and reputational damage. Failure to keep up with reporting changes also results in missed opportunities. Rapid and accurate reporting accelerates time-to-insights and powers critical decision-making at hyper-market speed.

If you are a senior stakeholder involved in reporting on sustainability, financials, and/or compliance and risk management, read on to learn how data agility is critical to fulfilling reporting requirements.



The data problem

At the heart of successful enterprise reporting is data. The five most important attributes of data for reporting are that it is: accurate, timely, clear and understandable, detailed, and comprehensive.

Businesses have to create reports quickly and easily, regularly, incrementally, and ad-hoc. And for diverse reporting requests – whether externally for regulators, clients, and investors, or internally to measure progress against strategic goals. To do this, enterprises need the ability to retrieve and integrate continuous data on demand from numerous systems in multiple formats. In short, the traditional manual approach to reporting won't cut it anymore. Automated reporting is a business-critical need.

The problem for many enterprises is that automated reporting is a pipe dream. They are held back by messy, chaotic data ecosystems plagued by data integration friction which makes reporting a serious headache. When businesses can't access, connect, and manage data flow across all sources, they increase the risk of non-compliance and slow their decision-making processes.

To solve the data problem, businesses must be agile in a way they never had to be before. They need to build an **agile infrastructure for enterprise reporting**, fit for a modern organization.



What is an agile data infrastructure?

An agile data infrastructure helps organizations make sense of complexity and eliminate data integration friction. It's a malleable set-up fit for messy and fluid data ecosystems, including applications and services in the cloud, and legacy, on-premises, and mainframe assets, across both internal and external systems.

An agile data infrastructure delivers elasticity, scale, and deep observability through a central mission control. It enables secure “any-to-any” data pipelines to be created and updated dynamically and in real-time from hundreds of sources. The in-built flexibility allows businesses to integrate core enterprise data into reports for greater consistency, accuracy, and depth of reporting, along with external sources like partner and supplier data.

With an agile infrastructure, business users no longer need to “get in the queue” for the IT and data engineering team to support a new data gathering and reporting request. It makes enterprise data available via self-service to multiple people who might need it to create a report – for instance, when the same data might be used for financial reporting, risk analysis, or internal rollups.



What are the challenges organizations face in enterprise reporting?

Successfully fulfilling complex reporting requirements requires timely access to accurate and consistent data. But without an agile data infrastructure to power enterprise reporting, line of business and data leaders face many challenges in getting what they need and meeting their obligations.

It's hard to access data in legacy systems

Data silos and sprawl

Evolving and increasingly complex regulations and standards

Integrating data from internal and external data sources, including files

Moving data from legacy to modern data environments

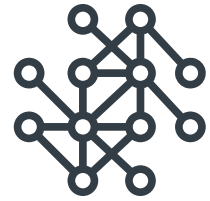
Heavy reliance on IT, lack of self-service access

Sharing data while adhering to regulatory requirements such as GDPR



How to build an agile data infrastructure for enterprise reporting in five stages?

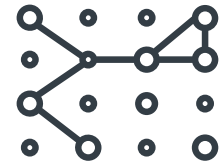
Stage 1



Establish data pipelines that are adaptable and resilient to change to ensure that pipeline breakage does not impact reporting accuracy

- Choose a data integration platform that provides a single, robust smart interface for managing all pipelines.
- Identify the pipelines that will do the heavy lifting of connecting data from source to destination for enterprise reporting. This can mean thousands of pipelines for the typical enterprise, including general ledger and billing systems, order management databases, supplier systems, and industry market data.
- Make sure the platform offers a breadth of connectivity that supports pipelines running in the cloud, on-premises, and in hybrid and multi-cloud environments.
- Select a platform that will automatically identify when infrastructure, schema or data changes occur so pipelines adapt to change without disruption. This will ensure resiliency is baked in and eliminate breakage. Pipeline breakage will result in reports containing out-of-date and inaccurate data.
- Create repeatable templates and pipeline fragments so brittle “one-off” hand-coded pipelines are no longer required. Pipelines should scale to thousands of different data sources by deploying common logic with minimal code to automate the creation of batch, Change Data Capture (CDC), and streaming pipelines.
- Ensure pipelines feed into a modern data storage destination like Snowflake or Databricks for self-service consumption. This means data engineers don’t have to manage and rebuild pipelines for every new reporting requirement.

Stage 2



Create a malleable infrastructure that's able to ingest and integrate data from any source in your ecosystem – including third parties, suppliers, vendors, and internal sources.

- Build an infrastructure that can connect to any source in your data ecosystem – not just internal systems, but also external sources like supplier and vendor files.
- This malleability guarantees there is no limit to transformation and gives businesses a rich palette to carry out customized and complex reporting requests using the latest data.
- Achieve this by using pre-built modern and legacy data source connectors and data processors so files and data can be easily ingested and incorporated into reports from:
 - **Operational systems** (OLTP, custom apps, mainframe)
 - **Internal sources** (mainframe, ERP apps)
 - **External sources** (partner and third-party data, supplier files)
 - **ESG sources** (carbon emissions, electricity consumption)
- Ensure data collection and ingestion engines are platform agnostic so data can be collated from any source to multiple platforms without rework, including hybrid and multi-cloud environments.

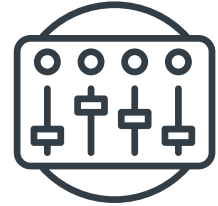
Stage 3



Embrace DataOps processes to support collaboration and create a skilled group of people who can pivot quickly – data integration is the lynchpin of DataOps.

- Data collection and ingestion must be operationalized to support stage one and two. This ensures that data can be integrated, transformed, and processed across multiple environments without friction.
- Achieve this by embracing DataOps – the practices and technologies that operationalize data integration for constant change and continuous delivery to power enterprise reporting.
- **There are three core principles of DataOps:**
 - **Design for change:** Build systems that are resilient to change and enable self-service for those who know the business and data needs best.
 - **Bring DevOps to data:** Enable automation and monitoring across the full lifecycle of productivity, from design to deployment and operations.
 - **Embrace emergent design:** Reveal data infrastructure, data structure, and semantics as they unfold instead of practicing design from the top down.
- Embedding DataOps principles will create a team of data leaders, analysts, and engineers who can tease order and discipline out of chaos to make enterprise reporting seamless and efficient. These are the people who can resolve data integration friction and keep insights flowing in real-time.
- DataOps also enables collaboration between central IT and line of business teams so data can be transported to destinations like Snowflake for collection. This allows users with the right permissions to access data to fulfill enterprise reporting tasks without roadblocks or delays in getting the data they need.

Stage 4



Enable a central mission control to provide complete observability and ensure data quality allowing the business to uncover data drift and guarantee accurate reports.

- The single, smart data integration platform of stage one becomes your mission control. It bridges the gap between new and legacy systems with a centralized hub for monitoring across environments.
- Use this mission control to deliver complete observability to central IT teams. Observability is crucial for showing how systems are connected and how data flows across the enterprise.
- Ensure there is deep visibility into data connections and flows across a hybrid landscape, including volume and throughput of data, as well as exactly what data is moving between components.
- Further support observability with data SLAs and rules that expose hidden friction to address issues in performance or data leakage before they affect reporting functions.
- Observability also allows teams to detect, control, and adjust to data drift – those unexpected, unannounced, and unending changes to data structure, semantics, and infrastructure that disrupt data flow.
- Employ auto-notifications based on user-defined triggers that observe pipelines for data quality, sizing, throughput performance, error rates, and private/sensitive information leakage so teams can respond quickly.

Stage 5



Embed security and don't cede control – protect the enterprise from one of today's most significant cyber threats posed by risky vendor and supplier data.

- There will always be some enterprise data that needs to be kept behind a firewall, and central IT teams must ensure they do not cede control of data internally or externally.
 - **Internally** – there is a risk when line of business teams act outside IT's purview and build their own data pipeline integrations that leave security gaps.
 - **Externally** – one of the most significant cyber threats to enterprises today comes from risky third-party data, such as compromised vendor and supplier data files. Security must be assured because these data sources are central to enterprise reporting.
- Alongside security vulnerabilities, a fragmented data supply chain also causes governance challenges, which is particularly relevant to regulated industries such as banking and financial services.
- The single data integration platform embeds security by allowing organizations to manage the full range of modern data sources required for enterprise reporting.
- Use pre-built connectors and destinations for common data sources and platforms so that data teams can monitor, secure, and control data movement to and from the full range of systems and data sources.
- Ensure policies and procedures governing how data is created, processed, and distributed are in place, granting access to data while complying with privacy and data safety laws and maintaining security.
- By establishing these guardrails, line of business users can access the data they need to complete enterprise reports without causing governance or security challenges.

Three key enterprise reporting use cases

Consistent reporting on business data is a critical requirement of compliance with a dynamic and continually evolving regulatory landscape. Iterative, incremental, and ongoing enterprise reporting has three key use cases.



USE CASE 1

Environmental, Social, and Governance (ESG) Reporting

Companies are expected to operate sustainably, take care of employees, and embed strong corporate governance. ESG reports summarize the qualitative and quantitative benefits of a company's ESG activities. This allows investors to screen investments, align investments to their values, and avoid companies with a risk of environmental damage, social missteps, or corruption. Third-party providers assign ESG scores based on reporting, including Bloomberg ESG Data Services, Sustainalytics ESG Risk Ratings, Dow Jones Sustainability Index Family, and RepRisk.

The EU currently has the most sophisticated set of ESG regulations. These ensure companies commit to ending net emissions of greenhouse gasses by 2050, decoupling economic growth from resources, and leaving no person or place behind. However, globally, there are also several notable ESG frameworks.

These frameworks are not yet mandatory but pertain to areas that will become so. [For example](#), The US Securities and Exchange

Commission has proposed mandatory climate disclosure for US-listed companies. China has signaled its intent to introduce ESG reporting rules, and the UK government has announced it will enforce mandatory carbon disclosure by 2025. Companies that can implement reporting processes for these areas now will ensure a smoother transition when compulsory regulation is rolled out.

Typical sources of ESG data:
Carbon Emissions, Utilities Data, HR Systems, Manufacturing Systems, Supplier Systems, Supplier Files

- Global Reporting Initiative (GRI)**
- The Sustainability Accounting Standards Board (SASB)**
- The Task Force on Climate-related Financial Disclosures (TCFD)**
- Carbon Disclosure Project (CDP)**
- Streamlined Energy and Carbon Reporting (SECR)**
- The Workforce Disclosure Initiative (WDI)**



USE CASE 2

Financial reporting

In all industries, decision-makers are required to complete regular and incremental financial reports, such as balance sheets, income statements, cash flow statements, and statements of shareholder equity. These reports power faster business decision-making, and serve and inform shareholders, creditors, investors, unions, employees, and regulators. The main objectives of financial reporting are to:

- Monitor income and expenses
- Ensure compliance
- Demonstrate robust governance
- Communicate essential data
- Support financial analysis and decision making
- Improve auditability and transparency

Regular financial reporting is also essential in projections of profitability, future industry position, and growth forecasting. Additionally, many of the reports that companies produce are made available for public review and are

monitored by government and private regulatory institutions to ensure fair trade, compensation, and compliant financial activities. It is therefore critical that the veracity and accuracy of financial reports can be ensured.

One of the key audiences for enterprises' financial reporting is regulatory authorities. This includes bodies like the SEC, IRS, and the Financial Accounting Standards Board (FASB) – which monitors GAAP (generally accepted accounting principles), the default accounting standard used by US companies. In the EU, all companies must prepare financial reports and statements in alignment with IFRS (international financial reporting standards).

Typical sources of financial data:

ERP, General Ledger, Billing Systems, Customer Order, Order Management Systems, Industry Market Data



USE CASE 3

Regulatory compliance

Industries including banking and financial services, manufacturing, healthcare, and transportation are heavily regulated. There is a long and continually evolving list of laws that organizations must follow to avoid non-compliance. The main objectives of regulatory compliance reporting are:

- Fulfilling regulatory requirements
- Proof of compliance
- Provide a synopsis for decision-makers
- Identify areas of improvement

Compliance reports prove an organization follows the stipulations of a particular regulatory standard. Violating laws, rules, and regulations can damage reputations, cause hefty penalties, and in some cases, even result in imprisonment or forced closure. These reports must be submitted to regulatory bodies under strict deadlines.

The burden of regulatory compliance is particularly heavy for banking, insurance, and financial services companies. Globally,

financial markets in individual jurisdictions have complex and stringent regulations. Businesses typically have to meet multiple reporting requirements for the different regions they operate in. For example:

- **Global:** FRTB (Fundamental Review of Trading Book), Lease accounting: IFRS 16 and ASC 842, BASEL III
- **US:** SOX (Sarbanes-Oxley), CAT (Consolidated Audit Trail), FORM PF (Private Fund)
- **EMEA:** GDPR, Solvency II, MiFID II, AIFMD, EMIR, PRIIPs
- **APAC:** HKMA (Hong Kong Monetary Authority)

Typical sources of compliance data:
ERP, General Ledger, Billing Systems, Customer Order, Order Management Systems, Industry Market Data

What advantages does an agile data infrastructure provide for enterprise reporting?

By following the five stages explained in this eBook, organizations can build an agile data infrastructure fit for the enterprise reporting needs of today's modern organizations.

With an agile data infrastructure, enterprises can:

- ✓ **Simplify data integration** for ESG, financial, and regulatory reporting through modern, resilient, and repeatable data integration pipelines.
- ✓ **Quickly and securely integrate and migrate data from internal and external sources, legacy systems, and files** into modern data platforms to accelerate ESG, financial, and regulatory reporting.
- ✓ **Easily and securely connect common ESG, financial, and regulatory data sources** and ingest data into modern data platforms.
- ✓ **Ensure a breadth of connectivity, modern data platform support, and flexibility** to run pipelines across on-prem, cloud, hybrid, and multi-cloud.
- ✓ **Eliminate data integration friction**, improving efficiency and scalability through repeatable, scalable data integration pipelines.
- ✓ **Bridge the gap between new and legacy systems** with a centralized control hub for data pipeline management and monitoring across environments.
- ✓ **Easily identify and manage data drift, including drift within files.**
- ✓ **Embed data SLAs and rules to expose hidden problems in data flows** to address performance, data leakage, quality, and other risks before they affect reporting veracity.
- ✓ **Gain access to pre-built data processors to power reporting.**



About StreamSets

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, data science, smart applications, and hybrid integration.